Национальная академия наук Беларуси
Объединенный институт проблем информатики

IV МЕЖДУНАРОДНЫЙ ФОРУМ IT-АКАДЕМГРАДА «ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В БЕЛАРУСИ»

13–14 октября 2025 года, Минск

ДОКЛАДЫ

Минск
ОИПИ НАН Беларуси
2025

УДК 004.9; 338.2; 656.13; 683.324

IV Международный форум IT-Академграда «Искусственный интеллект в Беларуси» : доклады, Минск, 13–14 октября 2025 г. – Минск : ОИПИ НАН Беларуси, 2025. – 380 с. – ISBN 978-985-7198-25-2.

Предлагаются доклады представителей государственных органов, учреждений образования, ученых и специалистов научно-исследовательских организаций, занимающихся исследованиями в области применения методов искусственного интеллекта.

Сборник предназначается для научных работников, преподавателей, аспирантов, студентов и широкого круга читателей.

Доклады, вошедшие в настоящий сборник, представляются в авторской редакции. Печатаются по решению программного комитета IV Международного форума IT-Академграда «Искусственный интеллект в Беларуси», а также редакционной коллегии государственного научного учреждения «Объединенный институт проблем информатики Национальной академии наук Беларуси».

Научные редакторы:
доктор военных наук, кандидат технических наук, доцент С. В. Кругликов;
кандидат технических наук, доцент С. Н. Касанин

ISBN 978-985-7198-25-2
© ОИПИ НАН Беларуси, 2025

Введение

Искусственный интеллект сегодня открывает человечеству все новые и новые возможности, помогая быстро и эффективно выполнять простые и сложные задачи в самых разных областях деятельности, таких как промышленность, сельское хозяйство, медицина, транспорт, экономика, государственное управление и многое другое. Сферы его применения продолжают расширяться и оказывают огромное влияние на бизнес, науку, образование, повседневную жизнь каждого.

С развитием облачных сервисов и увеличением вычислительных мощностей искусственный интеллект стал доступным для широкого круга пользователей. Благодаря искусственному интеллекту в образовании предоставляются возможности для создания индивидуальных образовательных программ и повышения эффективности преподавания, в сельском хозяйстве меняются традиционные подходы, по-новому решаются задачи продовольственной безопасности и развития, в промышленности на предприятиях сокращаются издержки и увеличивается продуктивность, в бизнесе и торговле производится анализ рыночных трендов и происходит быстрая адаптация к изменениям, прогнозируется сбыт и развивается уровень взаимосотрудничества. Алгоритмы искусственного интеллекта в медицине помогают врачам оперативно выявлять заболевания, ставить правильные диагнозы и повышать качество лечения. Внедрение его в транспортную сферу меняет представление о безопасности и скорости передвижений, а в интернете искусственный интеллект упрощает доступ к нужной информации и создает новый уровень взаимодействия с цифровой средой.

При исследовании проблем, касающихся искусственного интеллекта, решаются многие ключевые вопросы, связанные с развитием научной мысли, воздействием достижений в области вычислительных машин и робототехники на жизнь последующих поколений.

Государственное научное учреждение «Объединенный институт проблем информатики Национальной академии наук Беларуси» ежегодно с 2022 г. проводит Международный форум IT-Академграда «Искусственный интеллект в Беларуси», целями которого являются развитие и координация сотрудничества ученых и специалистов научно-исследовательских организаций, учреждений образования, занимающихся исследованиями методов искусственного интеллекта для поиска эффективных решений и практических рекомендаций по их применению в различных сферах жизнедеятельности. В задачи форума входят:

– ознакомление с современными разработками в области искусственного интеллекта и робототехники в рамках выставки форума «Искусственный интеллект в Беларуси»;

– формирование предложений по развитию взаимодействия научно-исследовательских организаций, учреждений образования в интересах продвижения и повышения мирового рейтинга белорусской школы искусственного интеллекта;

– развитие инфраструктуры поддержки научно-исследовательских организаций и образовательных учреждений, способствующей привлечению инвестиций в сферу разработки и практического применения методов искусственного интеллекта.

Сборник содержит доклады, представленные на IV Международный форум IT-Академграда «Искусственный интеллект в Беларуси», который проходил в Минске 13–14 октября 2025 г., по следующим направлениям:

1) регулирование в сфере искусственного интеллекта;

2) актуальные вопросы развития искусственного интеллекта в Республике Беларусь;

3) практические аспекты внедрения технологий искусственного интеллекта;

4) искусственный интеллект в медицине: сферы, технологии и перспективы, защита информации;

5) проблемные вопросы и пути совершенствования подготовки специалистов в сфере искусственного интеллекта.

В работе форума приняли участие представители государственных органов, учреждений образования, специалисты научных институтов и медицинских центров из Беларуси, России и Китая.

УДК 347

К вопросу о разработке понятия искусственного интеллекта в целях правового регулирования

Е. Н. Абрамова
Санкт-Петербургский государственный
экономический университет, Россия

Введение

Активное развитие технологий искусственного интеллекта привело к возникновению новых общественных отношений, породивших новые вызовы как для правовой науки, так и для законодателя. Одной из таких актуальных задач стало формирование понятийно-категориального аппарата, вошедшего в правовой обиход благодаря достижениям научно-технического прогресса. Однако современное положение дел демонстрирует существенное отставание правовой науки как от современных реалий гражданского оборота, так и от передовых технических научных разработок. Сегодня законодательство многих стран пытается нивелировать такой пробел и быстрыми темпами формирует правовое регулирование новых отношений, в том числе связанных с применением технологий искусственного интеллекта, однако подобная спешка при отсутствии фундаментальных научных разработок не всегда приводит к появлению качественного правового регулирования. В частности, одной из таких проблем является не всегда качественная проработка новых вводимых в оборот понятий и их дефиниций.

Понятийно-категориальный аппарат имеет особое значение при создании нормативной базы для совершенно новых, непривычных для традиционных правовых конструкций отношений, сопряженных с использованием неизвестных правовой науке понятий и инструментов. В отношениях, связанных с использованием технологий искусственного интеллекта, центральным выступает понятие самого искусственного интеллекта. Несмотря на то что его определения неоднократно приводились в том числе в нормативных актах, нельзя сказать, что в российском праве существует единообразное общепризнанное и тщательно разработанное понятие искусственного интеллекта.

Проблема неюридического характера легальных дефиниций

Одним из требований, которым должна отвечать легальная дефиниция, является соответствие определения назначению – способствовать

регламентации общественных отношений. Это означает, что определение должно быть направлено на установление юридической сути определяемого явления и быть практически применимым субъектами права.

Анализ научных и нормативных дефиниций искусственного интеллекта не позволяет сделать вывод о полном соответствии указанным требованиям предлагаемых определений. Так, в Российской Федерации в 2019 г. в п. 5 «Национальной стратегии развития искусственного интеллекта на период до 2030 года» [1], а позже в 2020 г. в подп. 2) ст. 1 Федерального закона «О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве» [2] искусственный интеллект был определен как комплекс технологических решений, который включает в себя информационно-коммуникационную инфраструктуру, программное обеспечение, процессы и сервисы. Несмотря на то что подобная формулировка прижилась в обиходе и была воспринята, в том числе в модельном законотворчестве стран СНГ, юридический смысл такого определения выявить затруднительно.

Во-первых, отсутствует согласованность данного понятия с определением технологического решения, предусмотренным в том же пункте Национальной стратегии в подп. м), где оно определяется как технология, программа для ЭВМ, база данных или их совокупность. Очевидно, что такое понятие гораздо уже тех составляющих искусственный интеллект как комплекс решений элементов, которые перечислены в определении искусственного интеллекта, хотя последнее представлено как частный случай технологических решений. Таким образом, невозможно установить, какое из двух указанных понятий является более широким, а какое – более частным.

Во-вторых, в определении искусственного интеллекта присутствуют не имеющие юридического наполнения слова, несущие исключительно бытовой смысл, например программное обеспечение, процесс и т. п. Их употребление не способствует однозначности и точности толкования, поскольку они понимаются разными субъектами исходя из их жизненного опыта, что приводит к домысливанию, додумыванию и, таким образом, не исключает искажение воли законодателя.

В-третьих, невозможно установить правовые последствия, которые могли бы вытекать из такого определения, поскольку предложенная формулировка не несет юридического наполнения, так как не позволяет дать правовую квалификацию явлению, установить элементы правоотношения или выполнить другую задачу, необходимую для правоприменения.

Как представляется, именно это и является наиболее важным для юридической дефиниции. Решения, в том числе технологические, могут как быть, так и не быть результатами интеллектуальной деятельности, а последние могут быть как охраняемыми, так и неохраняемыми, поэтому само по себе такое определение не позволяет ни выявить правовую природу, ни установить правовой режим. Также не ясно, выступает ли такой «комплекс» отдельным новым объектом гражданских прав. Из буквального толкования нормы следует, что это лишь совокупность отдельных объектов, не составляющих единый новый объект, поэтому не ясна цель его определения и возможность установления правового режима при отсутствии объекта гражданских прав.

В средствах массовой информации упоминается и совершенно новое определение искусственного интеллекта – как «любой системы данных, программного обеспечения или аппаратного средства, обладающей способностью обрабатывать данные и информацию способом, напоминающим разумное поведение...» [3]. По словам разработчиков законопроекта, такое определение «отвечает всем запросам относительно нейросетей» [3]. Однако относительно правовых требований данное определение не выглядит бесспорным. Во-первых, к искусственному интеллекту в нем предлагается относить множество разных по своей правовой природе явлений: информацию (данные), охраняемые результаты интеллектуальной деятельности (программное обеспечение) и вещи (аппаратные средства), которые являются разными объектами гражданских прав, обладают разным правовым режимом и не могут быть объединены в единую правовую категорию – систему, все ее фрагменты будут подчинены разным правовым режимам. Во-вторых, обладание способностью – это свойство субъектов права, а не объектов, каковыми являются все фрагменты названной системы. В-третьих, вызывает сомнения возможность способа напоминать разумное поведение. Способ – это система последовательных действий, которая не содержит внутреннего качества разумности. Указанные и иные недостатки предлагаемых дефиниций свидетельствуют о необходимости комплексного научного подхода к разработке дефиниций основополагающих для новой отрасли правового регулирования понятий.

Особо следует отметить, что неюридический характер правовых многих дефиниций основан на использовании не юридического словаря, а технической терминологии. Это не способствует обеспечению таких признаков любой дефиниции, как определенность и точность, недопустимость додумываний и пр. Юридическое определение, предназначенное для использования прежде всего правоприменителями, а также субъектами гражданского оборота и обеспечивающими их деятельность юристами, должно опираться на свойственный правоприменению словарь, а не технические

термины. В юридической литературе отмечается, что в результате развития IT-технологий стало распространяться правовое регулирование «с использованием научного инструментария, чуждого и сложного для восприятия юридического сообщества. Речь идет о новых для цивилистики понятиях точных наук, суть которых стала искажаться юристами. В результате додумывания, домысливания в доктрине стали анализироваться вымышленные правовые конструкции и правоотношения в цифровой среде, которых на самом деле не было и не могло быть из-за отсутствия соответствующих законодательных предписаний, регулирующих фактические отношения в digital-среде» [4]. Подобная ситуация представляется недопустимой и требует пересмотра выбранного подхода к правовому регулированию, которое должно быть направлено на установление определенности и стабильности гражданского оборота, обеспечения единообразного применения правил поведения, что возможно только в условиях формирования однозначных и точных формулировок правил поведения и дефиниций, основанных на использовании юридической терминологии.

Проблема несистемного характера легальных дефиниций

Выше уже отмечалось, что даже в рамках одного нормативного акта трудно добиться согласованности между разными нормами и дефинициями. Еще хуже ситуация, когда речь идет о согласовании дефиниций, предусмотренных разными актами. Как справедливо отмечают исследователи, когда «каждый новый закон применяет свои собственные перечни терминов, это нарушает единство и целостность нормативной основы отрасли, а в итоге – правовой системы в целом» [5].

Нередко разрабатываемые дефиниции практически невозможно применить в системе с другими определениями, данными смежным явлениям в других правовых актах. Подобная проблема присутствует не только в рамках национального законодательства, но и в модельном законотворчестве стран СНГ. Так, в ст. 2 модельного закона «О цифровом пространстве» цифровые технологии определяются как «совокупность знаний, методов и инструментов, применяемых при осуществлении оборота цифровой информации». Несомненно, технологии искусственного интеллекта являются цифровыми. Однако попытка применения в системе названной дефиниции цифровой технологии и определения технологии искусственного интеллекта как «технологии, основанной на использовании искусственного интеллекта», предусмотренного в ст. 2 модельного закона «О развитии искусственного интеллекта», приводит к выводу об их несогласованности.

Применение указанных двух дефиниций в системе дает неприменимое с точки зрения как логики, так и права понятие технологии искусственного интеллекта как «знания, методы и инструменты, основанные на использо-

вании искусственного интеллекта». Знания, основанные на использовании искусственного интеллекта, не могут быть признаны сами по себе технологиями искусственного интеллекта, поскольку с помощью последнего они были получены, но не являются им самим. Подобные выводы свидетельствуют о несогласованности предусмотренных в разных модельных актах дефиниций. Подобная ситуация возникает в силу масштабного нормотворчества в данной сфере, характерного для современного периода, когда разные акты разрабатываются практически одновременно. При этом у исследователей отсутствует фундаментальная база, на которую они могли бы опереться, поскольку таковую им приходится создавать практически с нуля самостоятельно. Это говорит о значительном отставании правовой науки от нужд гражданского оборота и современной экономики, о необходимости ускорения проведения системных научных исследований, которые представляют собой запрос уже даже не сегодняшнего, а вчерашнего дня.

Проблема отсутствия определений смежных понятий

Отправным для любой цифровой технологии, в том числе для технологий искусственного интеллекта, является термин «технология», однако его понятие не раскрывается в необходимой мере ни в законодательстве, ни в современной науке, хотя активно используется. При этом в научной юридической литературе отмечается, что это понятие является множественным, созвучно с понятием «техника» и зачастую приводит к смешению указанных терминов и неверному толкованию правовых норм [6]. В бытовом, техническом и нормативном применении данный термин в разных актах, разных странах и в разное время получал совершенно различное наполнение и смысл, поэтому совершенно нет однозначности в том, что такое технология в целом и технология искусственного интеллекта в частности. В принимаемых актах термин «технология», как правило, не определяется, а, наоборот, используется как само собой разумеющийся и через него определяются другие новые категории, в том числе технологии искусственного интеллекта.

Исследователи отмечают необходимость формирования понятий многих смежных явлений, носящих базовый характер для сферы применения искусственного интеллекта, но не имеющих общепризнанного понятия: сквозные цифровые технологии, интеллектуальные технологии, квантовый интеллект, вычислительный интеллект и другие: «самым важным для выбора стратегии правового регулирования технологии искусственного интеллекта является то, что применяются и развиваются самые разные технологии, которые основаны или связаны с интеллектуальными системами, но не всегда подпадают под понятие классического искусственного интеллекта» [7].

Результаты исследования

Проведенный анализ юридических определений искусственного интеллекта, в том числе нормативных, не позволяет признать их соответствующими юридической технике и отвечающими требованиям, предъявляемым к юридическим дефинициям.

Юридическая терминология и юридические определения не призваны выполнять функции толковых словарей, они предназначены для выполнения конкретных задач, связанных с правоприменением и реализацией права. Поэтому прежде всего необходимо обоснование необходимости той или иной дефиниции. Целям объяснения явлений служат наука, правовое и технологическое просвещение населения и другие институты, но не легальный юридический словарь. Последний предназначен для установления правовой квалификации, определения правовых последствий, правового статуса, правового режима и т. п.

Имеющиеся сегодня дефиниции искусственного интеллекта не позволяют найти его место в системе объектов гражданских прав или установить правовую природу, поскольку к нему относят комплекс, т. е. набор, совокупность совершенно разных по своему правовому режиму объектов, каждый из которых представляет собой самостоятельный объект гражданских прав. Определение искусственного интеллекта как комплекса технологических решений, каждое из которых имеет разную правовую природу, порождает вопрос о том, составляет ли такой комплекс отдельный самостоятельный новый объект гражданских прав, имущественный комплекс, для которого требуется разработка нового отдельного правового режима.

При существующем положении вещей термин «комплекс решений» играет лишь словообразовательную, но не правовую функцию, поскольку обозначает не новый самостоятельный комплексный объект гражданских прав, а некий набор разных по своей правовой природе сущностей, к которым относятся и разные объекты гражданских прав (вещи, объекты авторских прав, объекты патентных прав, др.), и всевозможные неправовые явления (процессы, сервисы и другие незнакомые цивилистике феномены). Это размывает понятие искусственного интеллекта, делает его нечетким и позволяющим разнообразные толкования, зависящие от жизненного опыта и представлений толкователя и даже от характера его специализации и профессии. Так, чрезмерное использование технической терминологии в нормативных дефинициях затрудняет их понимание правоприменителем и субъектами гражданского оборота, не обладающими соответствующими техническими компетенциями.

Уточнение в дефиниции на функцию искусственного интеллекта (позволяет имитировать когнитивные функции человека и получать результаты, сравнимые с РИД) еще более размывает определяемый предмет, по-

скольку делает недостаточными для правовой квалификации объективные внешние признаки. Такое уточнение означает, что для правовой квалификации в качестве искусственного интеллекта недостаточно обосновать наличие комплекса технологических решений предусмотренного законом состава, необходимо также исследовать технические возможности такого комплекса и оценить результат их функционирования. Однако в цивилистике цель использования объекта гражданских прав, как правило, не имеет никакого значения. Например, вещь останется вещью, как бы ее не использовали и что бы в результате ее применения не получали. Объективация признаков для признания явления конкретным объектом гражданских прав является общим правилом, позволяющим при любых обстоятельствах давать соответствующую правовую квалификацию. Также представляется сомнительной практическая применимость определения, ставящая квалификацию в качестве искусственного интеллекта в зависимость от наличия одновременно двух признаков: имитации когнитивных функций человека и получения результатов, сравнимых с его РИД. Во-первых, трудно представить правовые способы определения наличия таких функций. Во-вторых, использование в формулировке соединительного союза «и» исключает из круга определяемого явления те технологические решения, которые хоть и имитируют когнитивные функции человека, но не получают результаты, сравнимые или превышающие РИД человека. Представляется, что создание объектов интеллектуальной собственности является лишь одним из многих продуктов, которые человечество может получить благодаря технологии искусственного интеллекта.

Заключение

Таким образом, приходится констатировать, что современный понятийный аппарат о технологиях искусственного интеллекта формируется сегодня в критических условиях, характеризующихся отсутствием фундаментальной научной базы, проработанных доктринальных определений и обоснований необходимости тех или иных дефиниций для эффективного правового регулирования. Понятия включаются в создаваемые акты бессистемно, по принципу «все, что новое и непонятное», а не согласно требованиям юридической техники и исходя из потребностей правоприменения. Представляется, что качество формулировок легальных определений является базисом для качественного правового регулирования, которое не сможет предложить адекватных механизмов и решений правовых проблем, не определив, что именно оно должно регулировать.

Поэтому, несмотря на значительные достижения в науке и широкое распространение в экономике и обществе технологий искусственного интеллекта, до сих пор актуальным представляется проведение научных ис-

следований в целях установления круга понятий, определения которым необходимо дать для целей правового регулирования. При этом формируемые дефиниции должны выполнять юридические функции установления правовой природы и правовых последствий, носить исключительно юридический характер и быть адресованы, прежде всего, правоприменителю, который сможет применить новые нормы, например, в суде, для решения правовых задач: установления правовых последствий, правовой квалификации, выбора правового режима, защиты прав субъектов и т. п. Юридический словарь не следует использовать для решения не свойственных ему задач просветительского, образовательного, объяснительного или иного неправового характера.

Таким образом, проблема обоснования и разработки юридического понятийного аппарата в сфере искусственного интеллекта остается вопросом «злобы дня» в сфере правового регулирования общественных отношений, связанных с применением технологий искусственного интеллекта.

Список использованных источников

1. Национальная Стратегия развития искусственного интеллекта на период до 2030 года, утв. Указом Президента РФ от 10 октября 2019 г. № 490. – URL: https://a-ai.ru/wp-content/uploads/2024/03/Национальная_стратегия_развития_ИИ_2024.pdf (дата обращения: 18.10.2025).

2. О проведении эксперимента по установлению специального регулирования в целях создания необходимых условий для разработки и внедрения технологий искусственного интеллекта в субъекте Российской Федерации – городе федерального значения Москве» : Федеральный закон Российской Федерации от 24 апр. 2020 г. № 123.

3. В Госдуме сформулировали определение искусственного интеллекта // РБК. – URL: https://www.rbc.ru/technology_and_media/04/09/2025/68b8eec99a79477a1859ed11 (дата обращения: 18.10.2025).

4. Аюшеева, И. З. Гражданский оборот в условиях цифровизации: проблемы и перспективы развития / И. З. Аюшеева, Л. Ю. Василевская. – М. : Проспект, 2025. – С. 27.

5. Бачило, И. Л. О понятийном аппарате законотворчества и правовой науки на примере информационного права РФ / И. Л. Бачило // Понятийный аппарат информационного права» (ИГП РАН, отв. ред. И. Л. Бачило, Э. В. Талапина, сост. и ред. А. А. Антопольский). – М. : Канон-Плюс, 2015.

6. Рожкова, М. А. Понятие «технология» в контексте российского права / М. А. Рожкова, О. В. Исаева // Хозяйство и право. – 2021. – № 3. – С. 21–29.

7. Харитонова, Ю. С. Технология искусственного интеллекта и право: вызовы современности / Ю. С. Харитонова, В. С. Савина // Вестник Пермского университета. Юридические науки. – 2020. – № 49. – С. 524–549.

УДК 004.932

Обнаружение дыма на видеопоследовательности с использованием искусственной нейронной сети архитектуры трансформер

Е. Р. Адамовский✉, **Р. П. Богуш**
Полоцкий государственный университет
имени Евфросинии Полоцкой, Новополоцк, Беларусь
E-mail: e.adamovsky@psu.by

Введение

В настоящее время основным инструментом в задаче обнаружения дыма на кадрах видеопоследовательности являются искусственные нейронные сети (ИНС), поскольку они обучаются самостоятельно выделять значимые признаки из входных данных и могут работать в широком диапазоне условий наблюдения дыма за счет способности к обобщению информации. Наиболее распространены следующие архитектуры: сверточные сети (Convolutional Neural Network, CNN) и решения на их основе, например семейство YOLO (You Look Only Once); подходы с применением механизма внимания [1], включая модели ViT и DETR. Общими проблемами ИНС являются низкая скорость обработки, зависимость эффективности от качества обучающих данных, их объема и стратегии обучения.

Архитектура типа «трансформер» часто применяется для задач компьютерного зрения и использует механизм внимания (attention), который обрабатывает информацию с помощью структуры «энкодер-декодер». Трансформеры создавались для работы с текстом, однако были адаптированы под задачу обнаружения объектов на изображениях и продемонстрировали высокую точность. Детектирование дыма сводится к обнаружению и локализации объектов класса «дым», а сложность заключается в его полупрозрачности, хаотично меняющейся форме и отсутствии четких границ, поэтому применение трансформеров для такой задачи может дать превосходные результаты по сравнению с моделями на основе CNN. Современные обзорные работы [2–4] подчеркивают рост влияния трансформеров в компьютерном зрении, но их применение к задаче обнаружения дыма на видеокадрах остается малоизученным. Таким образом, целью работы является использование архитектуры трансформер для обнаружения дыма на кадрах видеопоследовательности.

Анализ архитектур трансформеров

Ключевой архитектурой трансформер для локализации объектов на изображениях является DETR (Detection Transformer) [5]. Подавляющее количество трансформерных моделей для задач компьютерного зрения основаны на DETR и используют его исходный код [6]. DETR состоит из нескольких модулей (рис. 1): backbone (модель CNN семейства ResNet); слои энкодера и декодера на основе механизма внимания; два полносвязных модуля (Feed-Forward Neural Network, FNN), предсказывающие класс и координаты ограничивающих рамок для объектов, которые ранжируются венгерским алгоритмом [7] с целью поиска наилучшего соответствия предсказаний истинным данным. В качестве носителя информации используются объектные запросы (object queries) фиксированного размера, которые проходят через декодер и содержат в себе данные о местоположении и классах обнаруженных объектов. Для запросов, которые не были отнесены ни к одному из детектируемых классов, назначается специальная метка no object.

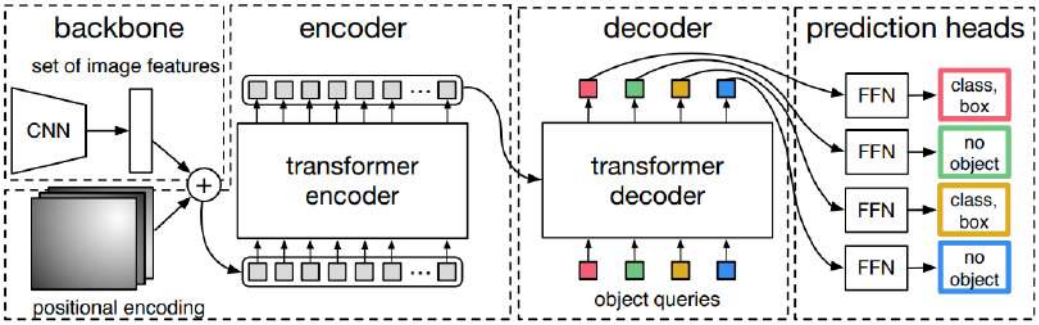


Рис. 1. Архитектура DETR [10]

Особенностью трансформеров является необходимость снабжения карт признаков позиционными данными (Spatial Encodings, SE), поскольку в дальнейшем они подвергаются преобразованию, при котором информация о взаимном расположении признаков теряется. В DETR используются фиксированные позиционные синусоидальные кодировки.

Медленная сходимость и низкая производительность DETR привели к созданию ряда модификаций, например [8, 9]. Стандартная модель DETR, предобученная на датасете COCO с 80 классами [10], использовалась для обнаружения дыма на отдельных изображениях в работе [11]. Для дообучения был собран набор данных из 13 000 кадров на основе данных сервиса AlertWildfire [12] размером 1920×1080 пикселей (Full HD), аугментированных случайными поворотами, отражениями, а также изменениями яркости и контраста. При этом исходная модель CNN ResNet-50 была заменена на одноклассовую и сброшены весовые коэффициенты объектных

запросов, а затем модифицированная архитектура Nemo-DETR дообучалась на новом датасете. Авторы [11] столкнулись с высокой частотой ложных срабатываний при обработке похожих на дым объектов, и в качестве решения использовали добавление в датасет «пустых» изображений без дыма и аннотаций, коллажей (аугментация «мозаика»), а также фиктивных рамок негативного класса с минимальным размером в случайных местах. В качестве дополнительного эксперимента объекты дыма были разделены по плотности на низкую, среднюю и высокую, но это привело к снижению точности обнаружения. В итоге Nemo-DETR достигла 42,3 % mAP (оригинальная модель на датасете COCO продемонстрировала 42 % mAP). При обнаружении дыма более информативными показателями являются AP_{50} и AR_{50}, их сочетание позволяет сбалансированно оценить склонность к пропускам и ложным срабатываниям при грубой локализации дыма, и для Nemo-DETR значения этих метрик равны 79 и 73,5 % соответственно.

В работе [13] использовалась модифицированная модель RT-DETR (Real Time DETR) с ResNet-18, в которой интегрированы модуль триплетного внимания, модуль ADown для снижения вычислительной сложности за счет масштабирования карт признаков и многоуровневая сеть слияния признаков. Для обучения RT-DETR был создан набор с отрицательными примерами: 12 564 изображения 640×640 пикселей, 10 154 с дымом и 2500 с другими объектами (например, белые и темные облака), которые аугментировались сдвигом, поворотом, отражением и мозаикой. Модель достигла точности локализации 96,1 % AP_{50} при скорости обработки 45 к/с.

Авторы [14] использовали улучшенную модель D-DETR, которая включала модули Multi-Scale Context Contrasted Local Feature и Dense Pyramid Pooling, улучшающие обнаружение объектов дыма с низким контрастом, что актуально для пожара в начальной стадии. С целью улучшения точности локализации применялся алгоритм NMS (Non-Maximum Suppression) и итеративное комбинирование рамок, которое объединяет прогнозы рамок с разных слоев декодера. При обучении применялся датасет из 10 250 изображений, размер пакета выбран равным 4. Точность модели составила 49,7 % mAP при скорости обработки 4 к/с на устройстве с GPU NVIDIA RTX 2080 Ti.

Архитектура RT-DETR [15] была подвергнута следующим изменениям: добавлен блок Coordinate Attention в слои модели ResNet для сохранения пространственной информации о положении объектов дыма; модифицирован многомасштабный энкодер; стандартная функция потерь DETR заменена на WShapeIoU для более эффективной обработки объектов дыма нестандартной формы. Был собран закрытый датасет из 16 316 изображений с разрешением 640×640 пикселей и преимущественно малыми дымовыми шлейфами. Модель обучалась в течение 250 эпох с размером пакета

32 на специализированном устройстве NVIDIA A100 с 80 ГБ оперативной памяти, за счет чего продемонстрировала скорость обработки 445 к/с. Значение метрики mAP составило 52,3 %.

Таким образом, приходим к заключению, что модели на основе DETR плохо обнаруживают объекты дыма, занимающие малую площадь в кадре, а точность их локализации может быть улучшена. Главной проблемой является склонность моделей к ложным обнаружениям. В рассмотренных работах используемые датасеты дополнительно включают в себя синтезированные данные. Применяется постобработка в виде алгоритма NMS. Аугментации включают поворот, отражение, изменение насыщенности, яркости и контраста, мозаику и сдвиг.

Сводная информация по характеристикам рассмотренных архитектур представлена в таблице.

Характеристики моделей на основе DETR
и параметры их обучения для обнаружения дыма

Модель	Состав датасета	Эпох	Размер пакета	Метрики, %					FPS	GPU
				mAP	AP_{50}	AP_S	AP_L	mAR		
Nemo-DETR [11]	13 000 кадр., FullHD	–	–	42,3	79,0	38,6	84,1	–	–	–
RT-DETR [13]	12 564 кадр., 640×640 пк	100	16	–	96,1	–	–	–	45	4×RTX 3080
RT-DETR-Smoke [15]	16 316 кадр., 640×640 пк	250	32	52,3	87,8	–	–	–	445	A100
D-DETR [14]	10 250 кадр.	50	4	49,7	88,4	36,9	62,3	60,1	4	RTX 2080 Ti

Из таблицы следует, что размеры используемых датасетов не превышают 15–20 тыс. кадров. Несмотря на то что оригинальная модель DETR обучалась в течение 500 эпох, для задачи детектирования дыма длительность обучения была сокращена. Данные о быстродействии противоречивы, поскольку для тестирования использовались GPU с различным уровнем вычислительной мощности. Тем не менее может быть сделан вывод о высоких требованиях DETR-подобных архитектур к производительности аппаратного обеспечения: в исследованиях использовались видеокарты NVIDIA, которые совместимы с технологией CUDA, не ниже 20-й серии, модели 80 или 90 с наибольшим объемом оперативной памяти. Точность детектирования рассмотренных решений сопоставима и составляет в метрике mAP 40–50 %, но при этом достигаются значения 80–95 % AP_{50}, т. е. модели обнаруживают дым в большинстве случаев, но не способны локализовать его достаточно эффективно.

Подготовка набора данных

Существующие датасеты с видеозаписями дыма, как правило, не содержат аннотаций и характеризуются низким разнообразием сцен. Ручная покадровая разметка имеет высокую трудоемкость и требует большого количества уникальных записей, поэтому с ее помощью может быть подготовлен лишь ограниченный объем данных. В качестве решения в некоторых исследованиях предлагается синтезировать сцены с помощью различных инструментов визуализации, что позволяет автоматизировать процесс разметки, но его эффективность ограничена визуальными отличиями синтезированного дыма от настоящего. Таким образом, решением является подготовка гибридного датасета, который содержит как записи реального дыма, так и синтезированные образцы, при этом значительная часть кадров не должна содержать никаких объектов дыма.

Итоговый датасет состоит из кадров, разделенных на тренировочные и валидационные, 18 760 и 2400 последовательностей соответственно, включая кадры без дыма и синтезированные данные, которые были сформированы и записаны на игровом движке Unity 5 (рис. 2), а затем интегрированы с помощью специального программного обеспечения, написанного на языке Python.

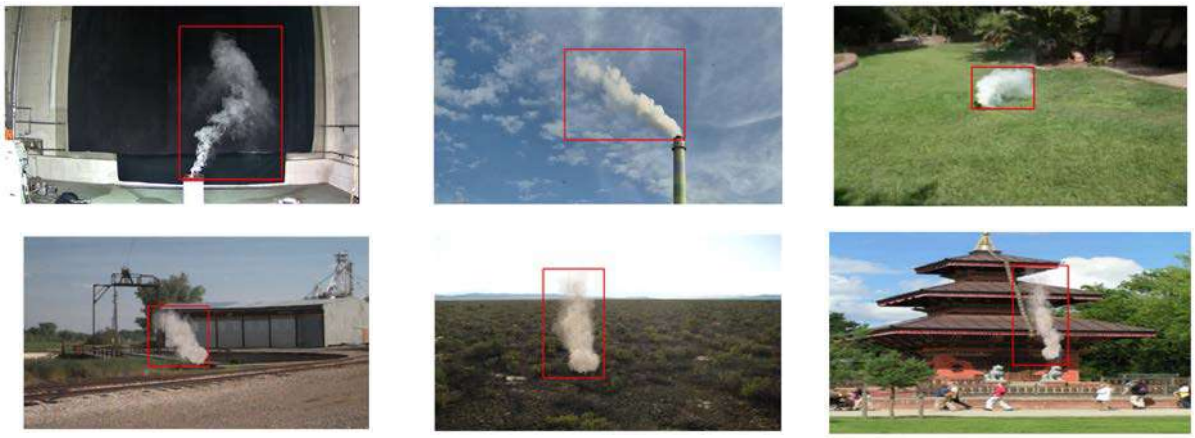


Рис. 2. Кадры с разметкой дыма из сформированного гибридного датасета (верхний ряд – реальный дым, нижний ряд – синтезированный дым)

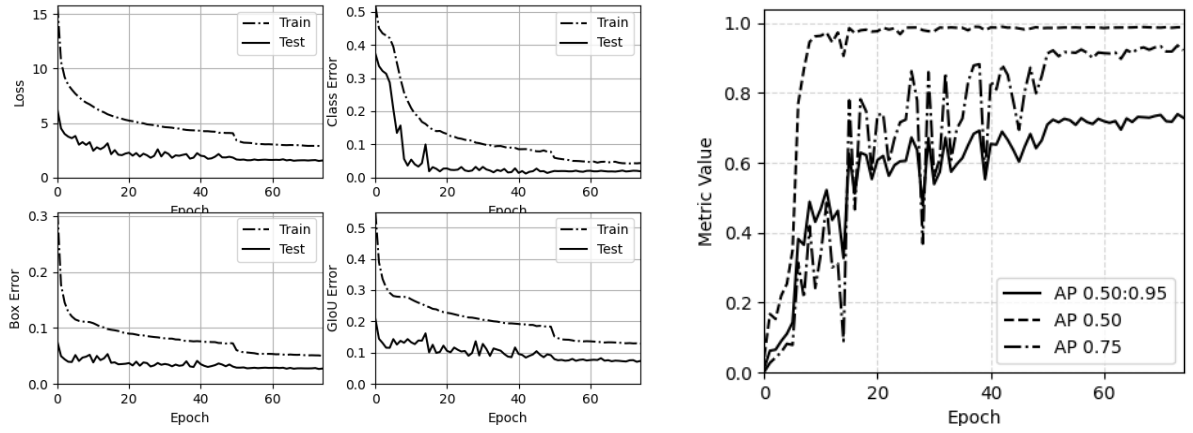
Результаты экспериментов

Модель была обучена со стандартными параметрами DETR, за исключением следующих значений: количество слоев энкодера и декодера снижено с 6 до 5, размерность каждого слоя и число головок внимания уменьшено в 2 раза, размер пакета увеличен с 2 до 6, длительность обучения составила 75 эпох, причем на 50-й эпохе скорость обучения была

автоматически снижена в 10 раз. В качестве CNN для извлечения признаков использовалась обученная на наборе ImageNet [16] модель ResNet-50.

Тестирование производилось на устройстве с GPU NVIDIA 2080 Ti. Скорость обработки одного изображения 1280×720 пикселей составила порядка 300 мс, что приблизительно соответствует кадровой частоте 3 к/с. Низкое быстродействие по сравнению с представленными в таблице решениями обусловлено использованием данных с более высоким разрешением и менее мощной задействованной видеокарты. Недостаток может быть скомпенсирован путем выборки из анализируемой последовательности кадров через заданный интервал времени.

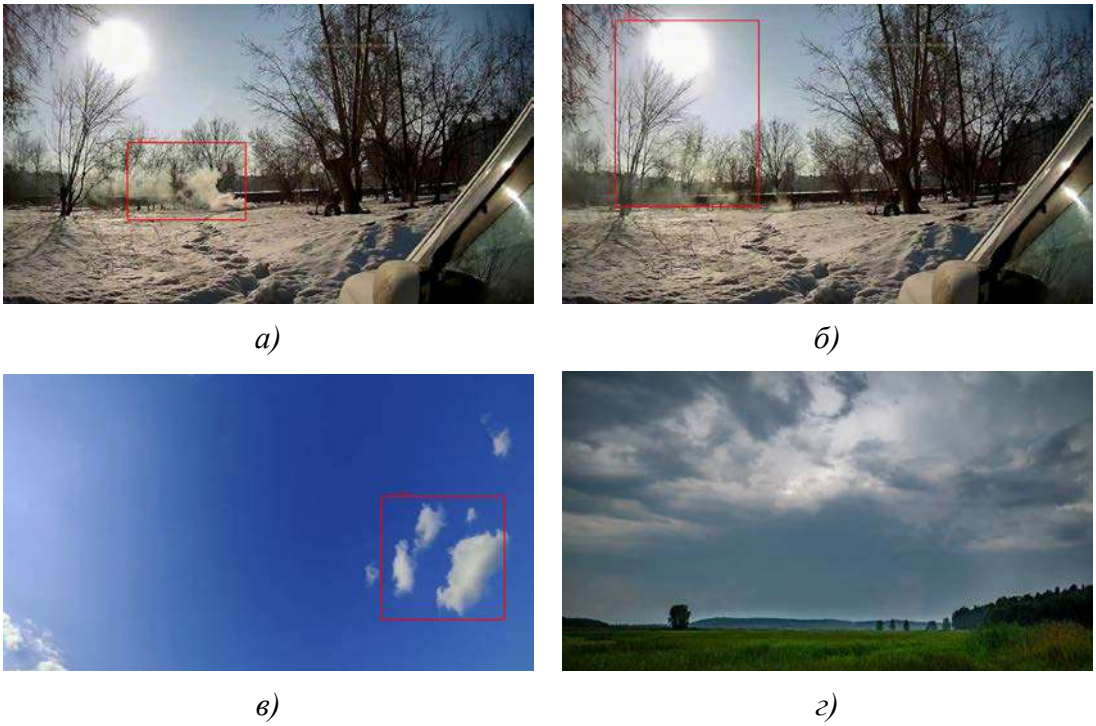
На рис. 3 представлены графики обучения модели: динамика изменения общей ошибки потерь и трех ее составляющих ошибок (рис. 3, *а*), а также метрика mAP и ее частные значения по различным порогам IoU (рис. 3, *б*). Из рис. 3 видно, что значения потерь на тренировочных данных значительно выше, чем на тестовых. Это объясняется тем, что к изображениям для проверки работы модели не применялись аугментации. В результате было достигнуто значение метрики 72,8 % mAP, причем значение AP_{50} составило 98,7 %, что говорит о высокой способности модели к обнаружению дыма при грубой локализации.



а) *б)*

Рис. 3. Графики обучения модели DETR: *а*) потери при обучении; *б*) метрика mAP

С другой стороны, модель продемонстрировала 79,7 % AR, т. е. в целом удовлетворительно справляется с поиском объектов дыма, занимающих значительную площадь в кадре, хотя и допускает пропуски. По полученным метрикам обученная модель превосходит ранее рассмотренные реализации на более чем 20 % mAP. На рис. 4 рассмотрены примеры работы модели с порогом уверенности 0,85.



а) *б)*

в) *г)*

Рис. 4. Примеры обнаружения дыма моделью DETR: *а*) корректное детектирование стелящегося дыма; *б*) ложное срабатывание на подсвеченную солнцем дымку; *в*) ложное срабатывание на облако; *г*) отсутствие ложных срабатываний на облака без четких контуров

Из рис. 4 видно, что модель способна уверенно обнаружить объект плотного дыма, но демонстрирует низкую точность при определении его границ. Дым с низкой плотностью или малой площади в кадре обнаруживается плохо. Также следует отметить проблему ложных срабатываний: детектор успешно определяет туман и плотные тучи как no object, но может неправильно идентифицировать кучевые облака как дым из-за высокой степени их визуальной схожести. При этом совершенно не похожие на дым объекты практически никогда не вызывают срабатывания.

Из анализа работы модели следует, что она способна обнаруживать дым с высокой точностью. Это подтверждается на примерах, отсутствующих в обучающей и тестовой выборках. Следовательно, архитектура трансформер может быть использована для детектирования объектов дыма. Однако эффективность была ограничена недостатками обучающего набора данных, который содержал недостаточно негативных примеров с дымоподобными объектами и позитивных примеров с рассеянным и значительно удаленным от точки наблюдения дымом.

Заключение

В данной работы была применена архитектура типа «трансформер» в виде модели DETR для обнаружения дыма на кадрах видеопоследовательности. Анализ существующих реализаций показал, что трансформерные архитектуры характеризуются низкой скоростью работы и склонны к ложным срабатываниям. Выявлены общие подходы к формированию обучающих наборов данных и их содержанию. Проведено обучение трансформерной модели DETR, описаны результаты экспериментов, которые подтверждают эффективность использования предложенной архитектуры для обнаружения дыма и направления на улучшение качества детектирования дыма.

Список использованных источников

1. Attention is all you need / A. Vaswani, N. Shazeer, N. Parmar [et al.] // Advances in Neural Information Processing Systems 30 : 31st Annual Conference on Neural Information Processing Systems (NIPS 2017) : Vol. 1 of 10. – Long Beach, California, USA, 4–9 Dec. 2017. – Curran Associates, Inc., 2018. – P. 5999–6009.

2. Transformers in Vision: A Survey / S. Khan, M. Naseer, M. Hayat [et al.] // arXiv preprint. – 2022. – arXiv: 2101.01169. – URL: https://arxiv.org/pdf/2101.01169 (date of access: 24.11.2025).

3. Image Segmentation Using Deep Learning: A Survey / S. Minaee, Y. Y. Boykov, F. Porikli [et al.] // IEEE Transactions on Pattern Analysis and Machine Intelligence. – 2021. – Vol. 44. – P. 3523–3542.

4. Object Detection With Deep Learning: A Review / Z. Q. Zhao, P. Zheng, S. T. Xu, X. Wu // IEEE Transactions on Neural Networks and Learning Systems. – 2019. – Vol. 30. – P. 3212–3232.

5. End-to-end object detection with transformers / N. Carion, F. Massa, G. Synnaeve [et al.] // European Conference on Computer Vision (ECCV). – Cham : Springer International Publishing, 2020. – P. 213–229.

6. DETR: End-to-End Object Detection with Transformers. – URL: https://github.com/facebookresearch/detr (date of access: 24.09.2025).

7. Du, K. An overview of object detection and tracking algorithms / K. Du, A. Bobkov // Engineering Proceedings. – 2023. – Vol. 33, № 1. – P. 22.

8. Deformable detr: Deformable transformers for end-to-end object detection / X. Zhu, W. Su, L. Lu [et al.] // arXiv preprint. – 2020. – arXiv:2010.04159. – URL: https://arxiv.org/abs/2010.04159 (date of access: 24.09.2025).

9. Dynamic DETR: End-to-End Object Detection with Dynamic Attention / X. Dai, Y. Chen, J. Yang [et al.] // Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV), 2021. – Curran Associates, Inc., 2022. – P. 2968–2977.

10. Microsoft COCO: Common Objects in Context. – URL: https://cocodataset.org (date of access: 24.09.2025).

11. Nemo: An Open-Source Transformer-Supercharged Benchmark for Fine-Grained Wildfire Smoke Detection / A. Yazdi, H. Qin, C. B. Jordan [et al.] // Remote Sensing. – 2022. – Vol. 14, № 16. – P. 3979.

12. AlertWildfire. – URL: https://www.alertwildfire.org (date of access: 24.09.2025).

13. Real-Time Smoke Detection in Surveillance Videos Using an Enhanced RT-DETR Framework with Triplet Attention and HS-FPN / L. Yang, Y. Cheng, F. Xu [et al.] // Fire. – 2024. – Vol. 7, № 9. – P. 387.

14. A small-target forest fire smoke detection model based on deformable transformer for end-to-end object detection / J. Huang, J. Zhou, H. Yang [et al.] // Forests. – 2023. – Vol. 14. – P. 162.

15. RT-DETR-Smoke: A Real-Time Transformer for Forest Smoke Detection / Z. Wang, L. Lei, T. Li [et al.] // Fire. – 2025. – Vol. 8, № 5. – P. 170.

16. ImageNet. – URL: https://image-net.org (date of access: 24.09.2025).

УДК 004.032.26

Нейросетевая модель профориентационной диагностики: разработка веб-приложения

К. В. Андренко✉, **В. А. Головко**,
С. А. Касперович, А. А. Крощенко
Брестский государственный
технический университет, Беларусь
E-mail: ksenia.andrenko2002@gmail.com

Введение

Сегодня особое внимание уделяется профориентации молодежи. Согласно Концепции развития профессиональной ориентации молодежи, утвержденной в 2014 г., профориентация направлена на оказание помощи молодым людям в осознанном выборе профессии, соответствующей их интересам, способностям и потребностям рынка труда [1]. Так, в 2024 г. был введен единый день профориентации в школах [2]. Среди мероприятий проводятся экскурсии на предприятия, встречи с профессорско-преподавательским составом вузов и колледжей, родителями, а также представителями профессий, дни открытых дверей в вузах и многое другое.

Недостаточная информированность молодежи о современных профессиях и требованиях, специальностях в вузах является одной из ключевых проблем. В 2024/2025 учебном году в школах активно внедряются новые формы профориентационной работы, такие как тренинги, экскурсии на предприятия и тестирование профессиональных склонностей. Например, онлайн-платформа «ПрофиТест», разработанная Республиканским институтом профессионального образования, уже помогла более 57 тыс. учащимся пройти диагностику профессиональных склонностей и интересов [3]. Эти меры направлены на устранение дисбаланса между спросом и предложением на рынке труда, который особенно остро ощущается в сельскохозяйственном и инженерном секторах, а также в образовании и медицине.

Для повышения качества результатов профориентации молодежи происходит внедрение современных технологий, хотя оно довольно ограничено. В 2024 г. Министерство образования Республики Беларусь подчеркнуло необходимость внедрения цифровых платформ и веб-сервисов для повышения доступности информации о профессиях, что приобретает особую важность в контексте конкуренции на рынке труда и перманентных перемен в требованиях к квалификации специалистов. Таким образом,

актуальность проблемы профориентации в Беларуси обусловлена необходимостью повышения осознанности выбора профессии, устранения дисбаланса на рынке труда и внедрения современных технологий в профориентационную работу.

Актуальность темы исследования связана с решением задач, поставленных такими программными документами, как Концепция развития профессиональной ориентации молодежи в Республике Беларусь; Национальная стратегия устойчивого социально-экономического развития Республики Беларусь на период до 2030 г.; Государственная программа «Образование и молодежная политика» на 2021–2025 гг.; Концептуальные подходы к развитию системы образования Республики Беларусь до 2020 г. и на перспективу до 2030 г.

Искусственный интеллект в профориентации

Применение искусственного интеллекта (ИИ) и машинного обучения в профориентации позволяет повысить точность результатов, объективность оценки, скорость обработки больших массивов неструктурированных данных, выявлять скрытые паттерны. Так, в исследовании [4] рассматривается применение машинного обучения в системах поддержки принятия решений на этапе постдипломного образовательного и карьерного консультирования. Преобладающие методы – обучение с учителем (supervised learning) и смешанное (mixed-method). Большинство исследований направлены на разработку экспертных систем и прогностических моделей, достигающих высоких показателей точности (precision) и верности (accuracy). Однако отмечаются и проблемы: непрозрачность (black-box problem), низкая объяснимость (explainability), предвзятость/смещение (bias) данных, конфиденциальность (privacy), прозрачность (transparency) работы системы.

Авторы [5] выделили основные роли ИИ в контексте профориентации: от инструментов-ассистентов до автономных систем под контролем человека. Среди предложений была отмечена необходимость развития data ecosystem – системы данных, включающей качественную и актуальную информацию. Например, учебные данные, данные об интересах студентов, рынке труда и т. д.

Несмотря на общие успехи, большинство существующих ИИ-решений не предлагают комплексной психологической характеристики абитуриента, ограничиваясь лишь рекомендацией специальности. Таким образом, актуальным становится создание профориентационного ассистента, который интегрировал бы данные о компетенциях вузов и психологические профили для повышения точности диагностики.

Профориентационный ассистент с использованием ИИ

Подготовка высококвалифицированных кадров во многом зависит от качественного набора студентов в учреждения высшего образования. Во-первых, группы, сформированные из мотивированных студентов, показывают более высокий уровень освоения учебного материала. Во-вторых, работа с вовлеченными студентами упрощает подготовку педагогов к занятиям, снижая необходимость тратить время на мотивацию студентов, позволяя сосредоточиться на освоении материала. Нужно подчеркнуть, что качественный набор предполагает и прямое решение проблемы дисбаланса на рынке труда, поскольку он обеспечивает приток мотивированных и компетентных кадров в стратегически важные секторы экономики.

Диагностирование в школьном возрасте предрасположенности к какой-либо дисциплине может упростить абитуриенту выбор специальности, подготовку к вступительным экзаменам и информировать родителей необходимой информацией. Также это может быть полезно для школ при планировании факультативных занятий. Для более точного прогнозирования лучше учитывать не только склонности к определенным предметным областям, но и психологические особенности абитуриента.

Традиционно системы, используемые для профориентации, сосредотачиваются на анализе предметных склонностей, реальных компетенций, игнорируя психологический профиль. Как показывает практика, потребность в общении, умение работать с монотонными задачами, стрессоустойчивость и другое напрямую связаны с успешностью закрепления в профессии в будущем или, наоборот, с профессиональным выгоранием. Например, склонность к изучению иностранных языков может указывать на разные специальности: переводчик, бизнес-аналитик (требуют постоянной динамичной коммуникации с людьми), лингвист и редактор (аналитическая и структурированная работа). Именно этот комплексный подход заложен в основу авторского MVP.

Для повышения точности диагностики предлагается MVP веб-приложения с применением ИИ. В качестве основных архитектур используются сверточные нейросети (CNNs) и большие языковые модели (LLMs). CNNs применяются для автоматизированного распознавания и классификации примитивов, необходимых для психогеометрического теста экспресс-диагностики типа личности (Стюзен Деллингер), обеспечивая высокую точность. LLMs, в свою очередь, задействованы для генерации отчета и обоснования рекомендаций, что повышает объяснимость системы. Предлагаемые результаты касаются не только рекомендуемой специальности, но и психологического профиля респондента.

Авторами были собраны таблицы компетенций по всем факультетам БрГТУ и БрГУ имени А. С. Пушкина в соответствии с требованиями образовательного стандарта и на их основе разработан комплекс из 116 вопросов, ответы на которые оцениваются методом шкалирования. Три ответа формулируются в свободной форме и обрабатываются LLMs. Используемый формат вывода результата: специальность + характеристика личности. Предполагается, что данное решение позволит успешно интегрировать инструмент в образовательную практику, повысит качество профориентационной работы и поспособствует осознанному выбору профессии.

Разработка веб-приложения

Приложение построено на клиент-серверной архитектуре (рис. 1), которая обеспечивает строгое разделение между уровнями представления и логики, взаимодействующими исключительно через API.

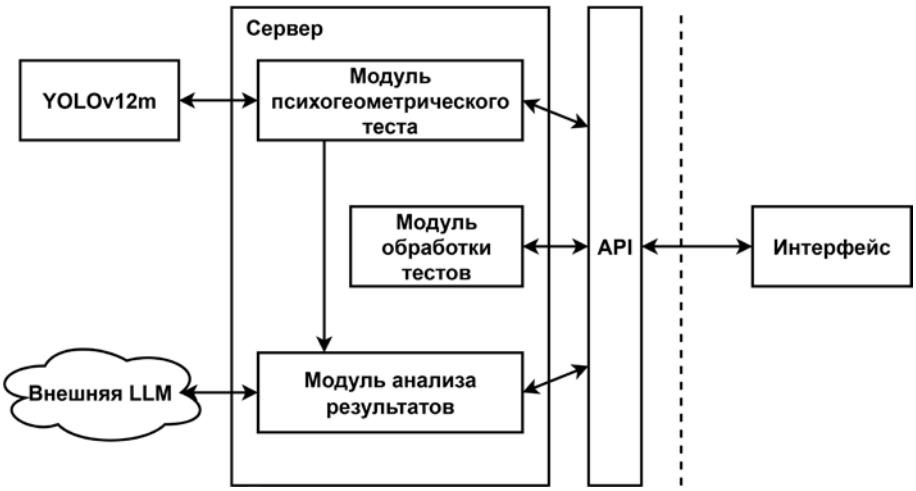


Рис. 1. Архитектура системы

Клиентская часть (фронтенд) отвечает за пользовательский интерфейс и интерактивность. Технологическая основа фронтенда – среда Node.js, которая используется для разработки интерфейсной части системы. Интерфейс предоставляет пользователю единое окно доступа ко всем функциям: формированию тестов, прохождению психогеометрического теста и просмотру результатов. Фронтенд отвечает за рендеринг, интерактивность и, что важно, за отправку пользовательских данных (изображения для анализа и ответы на тестовые вопросы) на сервер. На рис. 2 изображен главный экран приложения, на рис. 3 – пример одного из предлагаемых тестов.

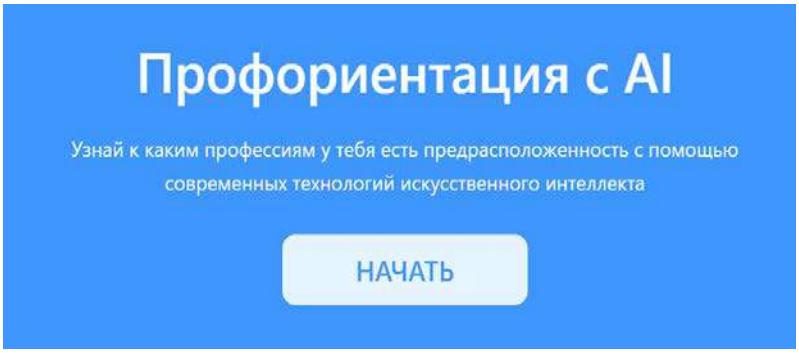


Рис. 2. Главный экран приложения

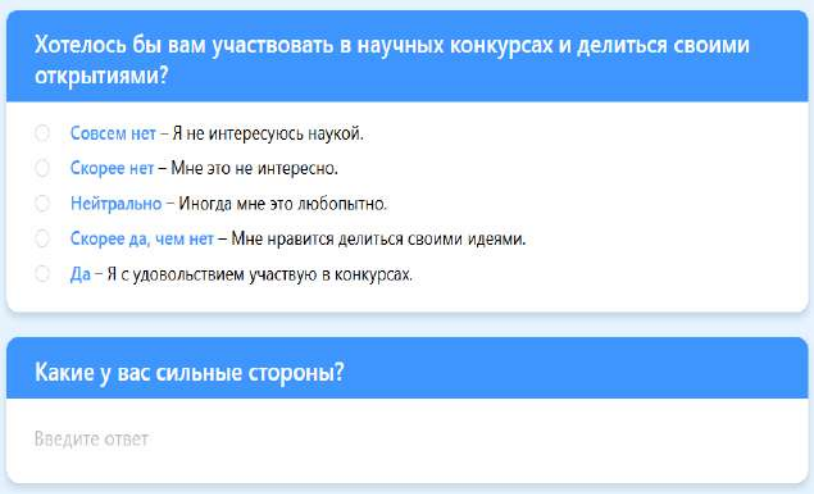


Рис. 3. Пример предлагаемых тестов с выбором и свободной формулировкой

Серверная часть (бэкенд) является ядром приложения, содержащим основную бизнес-логику и интеграцию с моделями ИИ. Сервер реализован на Python с использованием легковесного фреймворка Flask, что позволяет эффективно управлять запросами и предоставлять API-эндпоинты. Серверная часть включает три ключевых функциональных модуля.

1. Модуль психогеометрического теста. Основное назначение – распознавание геометрических примитивов (квадрата, круга, треугольника и т. д.) на изображениях, созданных и загруженных пользователем. В его основе лежит нейросетевая модель YOLOv12m [6], обученная решению задачи обнаружения (детекции) объектов. Модуль принимает изображение от клиента, обрабатывает его с помощью YOLOv12 для получения координат и классов обнаруженных фигур, а затем возвращает результат распознавания для дальнейшего анализа. Для обучения нейросетевого детектора использовалась подвыборка собранной базы рисунков из 71 изображения. Малый объем выборки частично компенсируется использованием предобученной архитектуры, снижающей требования к объему данных, а также большим количеством примитивов на каждом изображении. Тестирование

проводилось на 10 рисунках. Графики изменения ошибки в процессе обучения приведены на рис. 4, а результаты тестирования на отдельных изображениях – на рис. 5.

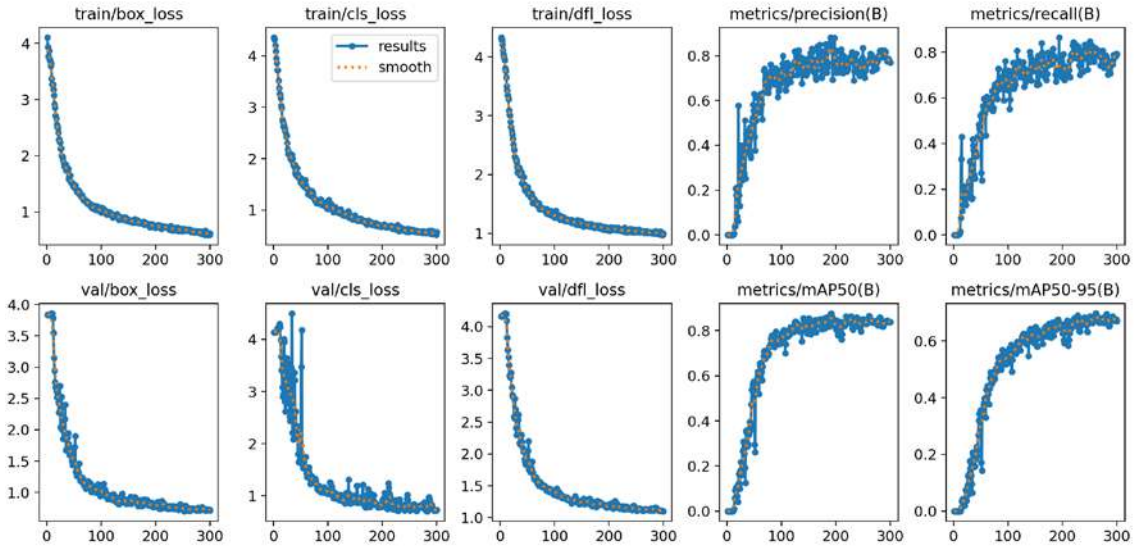


Рис. 4. Эволюция ошибки в процессе обучения

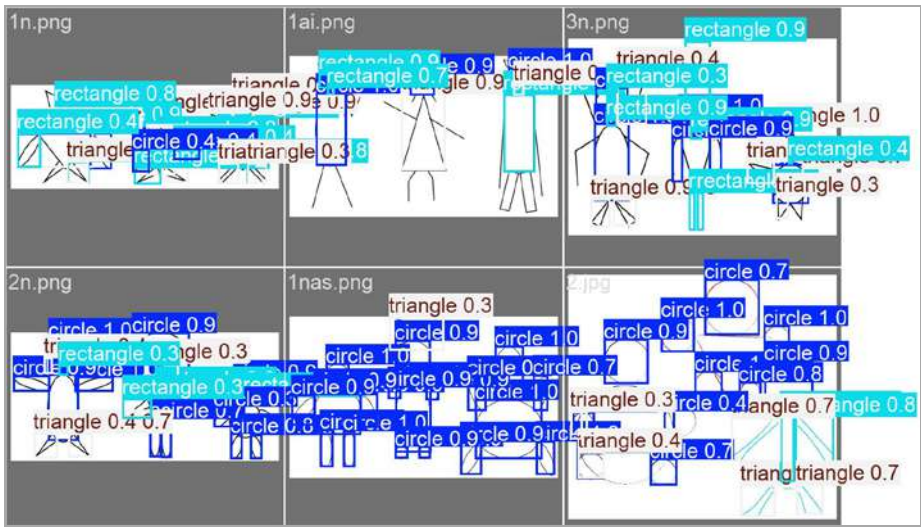


Рис. 5. Результаты тестирования модуля распознавания рисунков

Приведенные результаты могут быть существенно улучшены доразметкой всего объема имеющихся данных и дополнительными раундами обучения.

2. Модуль обработки тестов. Включает простую логику добавления, удаления и модификации базы тестов с настройками отображения.

3. Модуль анализа результатов. Назначение – комплексная обработка и интерпретация результатов всех тестов, включая данные, полученные из

психогеометрического теста. Этот процесс включает формирование промптов – создание структурированных запросов на основе собранных данных. Затем эти промпты отправляются через вызовы API к внешней большой языковой модели, такой как OpenAI GPT, для получения глубокой, развернутой интерпретации и персонализированного анализа. Визуализация работы данного модуля представлена на рис. 6.

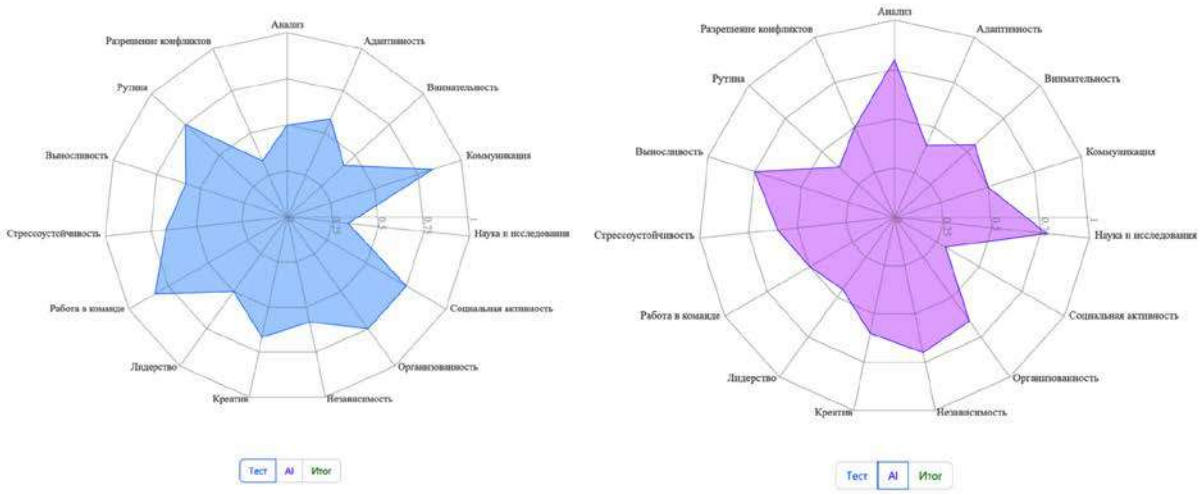


Рис. 6. Результаты теста и ИИ

Таким образом, стек используемых технологий выглядит следующим образом: клиентская часть/фронтенд использует Node.js для разработки и запуска интерфейса; серверная часть/бэкенд использует Flask (Python) для предоставления API, обработки запросов и управления бизнес-логикой. Компьютерное зрение реализовано с помощью YOLOv12 для классификации примитивов, а ИИ – через API большой языковой модели для анализа и интерпретации результатов. Важно отметить, что модель YOLOv12 была обучена на кастомном датасете для достижения высокой точности и скорости распознавания в реальном времени (Real-Time Detection), что является критичным для психогеометрического теста.

Заключение

Одними из стратегически важных направлений в Республике Беларусь являются качество подготовки кадров и баланс на рынке труда, что напрямую связано с профориентацией молодежи. Традиционные формы профориентационной работы способствуют повышению информированности о профессиях и специальностях, по которым может происходить обучение. Необходимым условием эффективности является автоматизация процесса при использовании технологий ИИ и машинного обучения.

Анализ существующих исследований показал отсутствие комплексного подхода к профориентации с фокусом на предметных склонностях, что потенциально может привести к различным рискам (профессиональное выгорание, потеря мотивации, смена сферы деятельности).

В качестве одного из вариантов решения данной проблемы, авторами предлагается MVP веб-приложения, интегрирующее технологии ИИ: CNNs и LLMs.

Предложенный MVP является шагом в направлении интеграции передовых технологий в образовательную практику. Его внедрение может способствовать более осознанному выбору профессии школьниками, повышению качества набора студентов в вузы, а также снижению дисбаланса между спросом и предложением на рынке труда. В перспективе подобные решения станут неотъемлемой частью системы профориентации, формируя основу для устойчивого социально-экономического развития Республики Беларусь.

Список использованных источников

1. Концепция развития профессиональной ориентации молодежи в Республике Беларусь // Министерство образования Республики Беларусь. – URL: https://edu.gov.by/en-uk/urovni-obrazovaniya/srednee-spetsialnoe-obrazovanie/upravlenie-professionalno-tekhnicheskogo-i-srednego-spetsialnogo-obrazovaniya/normativnye-pravovye-dokumenty/Концепция%20развития%20профессиональной%20ориентации%20молодежи%20в%20республике%20Беларусьпрофориентац.%20утвержд..docx (дата обращения: 15.01.2025).

2. В белорусских школах введут единый день профориентации // БЕЛТА. – URL: https://belta.by/society/view/v-belorusskih-shkolah-vvedut-edinyj-den-proforientatsii-665568-2024/ (дата обращения: 05.10.2024).

3. Комплекс онлайн-методик «ПрофиТест» [сайт] // РИПО. – URL: https://profitest.ripo.by/public/main (дата обращения: 18.09.2024).

4. Manganello, E. Machine learning for post-diploma educational and career guidance: a scoping review in AI-driven decision support systems / E. Manganello, E. Rasca, A. Villa [et al.] // Frontiers in Education. – 2025. – Vol. 10. – DOI:10.3389/feduc.2025.1578979.

5. Artificial Intelligence for Career Guidance – Current Requirements and Prospects for the Future / S. Westman, J. Kauttonen, A. Klemetti, N. Korhonen // International Journal of Emotional Education. – 2021. – Vol. 9, iss. 4. – DOI: 10.22492/ije.9.4.03.

6. YOLO12: детекция объектов с акцентом на внимание [сайт]. – URL: https://docs.ultralytics.com/ru/models/yolo12/ (дата обращения: 07.09.2025).

УДК 004.8 (330.34; 338.2)

Стратегии формирования национальных экосистем ИИ: сравнительный анализ ведущих стран мира

Э. Бабоян, Я. А. Дявго✉, А. С. Качан
Институт экономики НАН Беларуси,
Минск
E-mail: yana.dyavgoo@gmail.com

Введение

Искусственный интеллект (ИИ) эволюционировал от узкоспециализированной области компьютерных наук до ключевого драйвера глобальных экономических и социальных преобразований, которые характеризуют наступление Четвертой промышленной революции. Его интеграция в различные отрасли коренным образом меняет модели производства, управления и потребления, определяя контуры будущей мировой экономики. В этом контексте актуальность темы исследования обусловлена тем, что уровень развития и внедрения технологий ИИ становится прямым индикатором конкурентоспособности государства на мировой арене, непосредственно воздействуя на его экономический рост, технологический суверенитет и геополитическое влияние.

Ожидается, что ИИ внесет значительный вклад в глобальный ВВП, а страны – лидеры в этой области получат непропорционально высокие выгоды [1]. Все это порождает острую глобальную конкуренцию за лидерство в сфере ИИ, которая приобретает характер новой «технологической гонки». Ведущие экономики мира, осознавая стратегические риски отставания, активно разрабатывают и реализуют национальные стратегии, направленные на ускоренное формирование целостных и жизнеспособных экосистем ИИ. Однако подходы к их построению существенно различаются: от жестко централизованной, государственно-управляемой модели до децентрализованной, рыночно-ориентированной. Это создает научную и практическую проблему, заключающуюся в отсутствии систематизированного сравнительного анализа данных стратегий, который позволил бы выявить их сильные и слабые стороны, общие тенденции и уникальные национальные особенности.

Теоретико-методологические основы формирования экосистемы ИИ

В данном исследовании под национальной экосистемой ИИ понимается сложная, динамичная и взаимосвязанная система формальных и неформальных институтов, норм, ресурсов и субъектов, которые совместно участвуют в создании, распространении и применении технологий ИИ на территории государства. Эта экосистема функционирует на стыке государственной политики, академической науки, бизнеса и общества, а ее эффективность определяет способность страны занять лидирующие позиции в глобальной конкуренции за технологическое лидерство.

Можно выделить пять ключевых структурных компонентов национальной экосистемы ИИ (рис. 1): *государство* – стратегический координатор; *наука* – генератор знаний; *инфраструктура* – обеспечение данными и вычислительными мощностями; *кадры* – ключевой человеческий ресурс; *бизнес* – сфера внедрения и практического применения. Устойчивое развитие экосистемы возможно только при условии активного взаимодействия и циркуляции ресурсов (знаний, данных, кадров, капитала) между всеми структурными элементами.

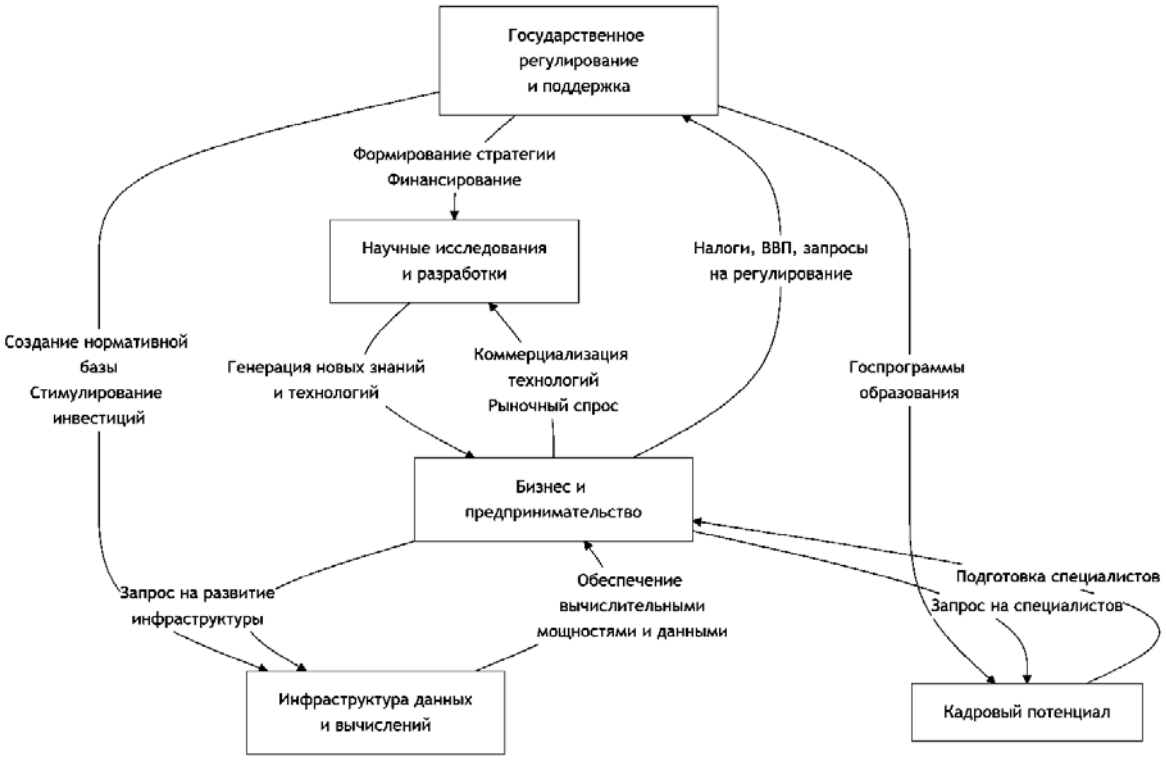


Рис. 1. Схема взаимодействия субъектов национальной экосистемы ИИ

В современных геополитических условиях формирование целостной национальной экосистемы ИИ становится не только вопросом экономиче-

ской конкурентоспособности, но и основой для обеспечения технологического суверенитета государства. Это, в свою очередь, позволяет:

– снизить зависимость от иностранных технологий;

– сфокусировать национальные ресурсы на прорывных направлениях, имеющих ключевое значение для безопасности и развития экономики;

– создать целостную экосистему, которая обеспечит устойчивость национальной экономики за счет развития собственного кадрового и исследовательского потенциала.

Анализ стратегий ведущих стран в сфере создания ИИ-экосистем

В условиях глобальной цифровой трансформации и стремительного развития технологий ИИ становится все более актуальной задача объективной оценки готовности стран к созданию и развитию национальных экосистем ИИ. Для этого международные организации и исследовательские центры разрабатывают специализированные индексы, позволяющие сопоставлять уровень технологического развития, институциональной поддержки, инновационного потенциала и качества цифровой инфраструктуры различных стран.

Среди наиболее авторитетных инструментов оценки можно выделить Government AI Readiness Index[1], отражающий готовность правительств к внедрению ИИ в управление и экономику; Global AI Index[2], который фиксирует глобальные тренды в инвестициях, публикационной активности и разработках; Global Innovation Index[3], оценивающий инновационную среду и результаты научно-технологического развития; Network Readiness Index[4], позволяющий оценить цифровую связность и инфраструктуру; а также E-Government Development Index[5], характеризующий степень цифровизации государственного сектора.

Анализ этих индексов по ряду стран – лидеров в сфере ИИ (США, Китай, Великобритания, Сингапур, Дания, Эстония), а также по государствам с формирующейся экосистемой (Беларусь, Россия) позволяет выявить сильные и слабые стороны национальных стратегий, определить их потенциал и ограничивающие факторы, а также построить более обоснованные рекомендации для развития ИИ в странах постсоветского пространства (рис. 2).

[1]Government AI Readiness Index 2024. – URL: https://oxfordinsights.com/ai-readiness/ai-readiness-index/ (date of access: 03.09.2025).
[2]The Global AI Index. – URL: https://www.tortoisemedia.com/data/global-ai#data (date of access: 03.09.2025).
[3]Глобальный инновационный индекс. – URL: https://www.wipo.int/ru/web/global-innovation-index (date of access: 03.09.2025).
[4]Network Readiness Index 2024. – URL: https://networkreadinessindex.org/ (date of access: 03.09.2025).
[5]E-Government Development Index (EGDI). – URL: https://data360.worldbank.org/en/dataset/UN_EGDI (date of access: 03.09.2025).

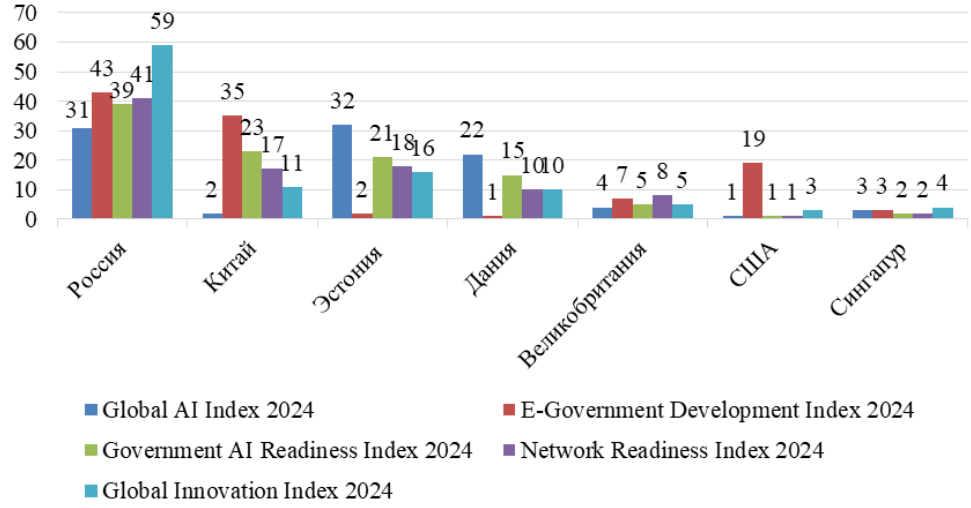


Рис. 2. Профили стран (выборочно) по основным индексам, ранг

Развитые экономики (США, Великобритания и др.) уверенно лидируют в цифровых и инновационных индексах, формируя основу глобальной экосистемы ИИ. Китай активно сокращает разрыв благодаря масштабным инвестициям и развитию прикладных решений. Россия остается в группе догоняющих, ИИ-экосистемы находятся на стадии становления.

Рассмотрим подробнее национальные стратегии создания и развития экосистем ИИ выбранных здесь стран (табл. 1). Малыми цифровыми лидерами в данной области являются Дания, Эстония и Сингапур.

Таблица 1

Сравнительная характеристика национальных стратегий создания и развития экосистем ИИ [2–4]

Страна	Институциональная стратегия	Финансовая стратегия	Кадровая стратегия
Дания	Ключевые документы: EU AI Act; Закон о дополнительных положениях к регламенту об искусственном интеллекте; Strategic Approach to AI. Ключевые институты: Агентство цифрового правительства; Датский орган по защите данных; Судебная администрация	Целевые вложения в Digital Taskforce, Centre for AI in Society, платформу датских языковых моделей и открытые датасеты; исследовательские резервы по этике и объяснимости; долгосрочная поддержка Pioneer Centre; инфраструктура EuroHPC и Gefion	Digital Taskforce on AI; Centre for AI in Society; партнерство с Pioneer Centre for AI и Danish Foundation Models; модули повышения квалификации для госслужащих и бизнеса

Окончание табл. 1

Страна	Институциональная стратегия	Финансовая стратегия	Кадровая стратегия
Эстония	Ключевые документы: EU AI Act; National AI Strategy 2.0. Ключевые институты: Мин-во экономики и коммуникаций; Мин-во образования и исследований; Мин-во юстиции	Мультифондовое финансирование через национальную программу «НИОКР и инновации»; гранты Horizon Europe; европейские структурные фонды; ваучеры Enterprise Estonia; конкурсы Эстонского исследовательского совета; налоговые льготы	Программа ProgeTiger; Elements of AI; магистратура Data Science; постдокторские гранты EIS; курсы и практические лаборатории EIS; инновационные хабы и индустриальные акселераторы
Сингапур	Ключевые документы: Model AI Governance Framework; Generative AI Playbook; National AI Strategy 1.0 и 2.0. Ключевые институты: SNDGG; Консультативный совет по этичному ИИ; PDPC; IMDA; GovTech; A*STAR; NRF; MAS	Централизованное финансирование через SNDGG и NRF программ AI Singapore (Grand Challenges; 100 Experiments; ConnectAI; LearnAI; AI Apprenticeship); бюджеты PDPC и IMDA на рамки управления ИИ и AI Verify; GovTech и A*STAR на GenAI Sandbox и Project Moonshot	Массовое образование (AI For Everyone; LearnAI; GenAI Sandbox хакатоны); гранты AI Singapore и стипендии NRF; курсы PDPC/IMDA и обучение AI Verify; междисциплинарные хабы SEA-LION и Project Moonshot

Сравнительный анализ стратегий Дании, Эстонии и Сингапура показывает, что все три страны исходят из необходимости институциональной поддержки и правового регулирования ИИ, но используют разные модели управления. Дания и Эстония опираются на общеевропейский «AI Act», но Дания выбрала централизованную схему с опорой на существующие ведомства, тогда как Эстония внедрила распределенную модель с межведомственным управлением. Сингапур развивает собственную нормативную базу и уникальные инструменты валидации (AI Verify), что делает его примером национальной автономии и технологической самодостаточности.

В части финансирования Дания акцентирует внимание на целевых инвестициях в исследовательские центры, языковые модели и инфраструктуру. Эстония использует мультифондовую систему с привлечением европейских средств и налоговых стимулов, а Сингапур демонстрирует максимально централизованный подход через государственные фонды и коор-

динирующие агентства. Интересным является кадровый аспект: Дания делает ставку на межведомственные рабочие группы и специализированное обучение чиновников, Эстония интегрирует ИИ-компетенции в школьное образование и развивает академическую подготовку через магистратуру и постдокторские программы, а Сингапур использует массовое образование и инновационные песочницы для широкого вовлечения населения и бизнеса.

Так, опыт трех стран иллюстрирует разные траектории формирования национальных экосистем ИИ. Каждая из них демонстрирует полезные практики: Дания – эффективное использование существующих институтов, Эстония – системное вовлечение образования и частного сектора, Сингапур – институционализированное сочетание этики, регулирования и масштабного обучения.

Перейдем к обзору стратегий развития экосистем ИИ в странах – глобальных инвесторах (вендорах) ИИ: США, Китае и Великобритании (табл. 2).

Таблица 2

Сравнительная характеристика национальных стратегий создания и развития экосистем ИИ [5–8]

Страна	Институциональная стратегия	Финансовая стратегия	Кадровая стратегия
США	Ключевые документы: Artificial Intelligence Roadmap; Программа AI Action Plan; Memorandum on Advancing the United States' Leadership in AI; Harnessing AI for National Security Objectives; Fostering Safety, Security and Trustworthiness. Ключевые институты: Office of Science and Technology Policy; National AI Initiative Office (NAIIO); Select Committee on Artificial Intelligence; Chief Digital and AI Office (CDAO) Artificial Intelligence IWG; National Institute of Standards and Technology (CAISI и AISIC); AI Research Institutes, под эгидой NSF	Программа финансовой помощи, включающая кредиты, гарантии по кредитам, гранты, налоговые послабления и соглашения о закупках; лояльные условия для получения земли под дата-центры и фабрики по производству полупроводников; государственное финансирование в рамках программы «Исследования и разработка сетевых и информационных технологий» (Networking and Information Technology Research and Development, NITRD)	Программы AI Research Institutes (NSF), National AI Research and Development Strategic Plan (развитие STEM-образования, расширение доступа к AI-курсам в университетах и школах); K–12 AI Education Initiatives; NSF Graduate Research Fellowships; DARPA финансирует образовательные модули в университетах в рамках своих исследовательских грантов

Окончание табл. 2

Страна	Институциональная стратегия	Финансовая стратегия	Кадровая стратегия
Китай	Ключевые документы: Программа «Искусственный интеллект+»; «Сделано в Китае 2025»; «Трехлетний план действий в области ИИ»; Программа развития ИИ нового поколения (дорожная карта до 2030 г.); Имплементация видения Шанхая о развитии ИИ нового поколения; Новые руководящие принципы для развития индустрии ИИ; Белая книга по стандартизации ИИ; Некоторые меры содействия развитию индустрии ИИ. Ключевые институты: Семь ведомств, в том числе на Cyberspace Administration of China (CAC), Министерство образования, Министерство науки и технологий, National Development and Reform Commission	Фонды (Национальный инвестиционный фонд ИИ, проект «Передача данных Восток-Запад», Пекинский AI-фонд, Шанхайский материнский фонд, провинциальные программы, напр., в Гуандуне, и др. имеющиеся ресурсы, координирование распределения капитала международных и внутренних инновационных институтов; снижение налоговой нагрузки; привлечение предприятий и общества для увеличения инвестиций и построение новой модели многопартийной поддержки с использованием партийного и социального капитала	AI Innovation Action Plan for Institutions of Higher Education; расширение приема на ИТ/инженерные специальности; создание национальных AI Research Institutes; классические программы привлечения «Thousand Talents» и ряд региональных инициатив/грантов для привлечения зарубежных специалистов и возвращения соотечественников; собственные академии, магистерские программы и стажировки больших корпораций; масштабирование и K–12
Великобритания	Ключевые документы: National AI strategy; AI Sector Deal; Defence Artificial Intelligence Strategy; White Paper A pro-innovation approach to AI regulation. Ключевые институты: Department for Science, Innovation & Technology; Office for Artificial Intelligence	Программа AI Growth Zones; AI Upskilling fund; снижение дефицита финансирования инновационных компаний; финансирование за счет средств государственного банка экономического развития British Business Bank, пенсионных фондов и грантовых программ в области цифровизации (Digital Growth Grant); имплементация в законодательство рекомендаций из обзора Hill Review для привлечения заемного капитала через фондовую биржу; реформа системы госзакупок – доступ к закупкам для малого инновационного бизнеса станет простым и гибким	Наличие ведущих мировых исследовательских институтов в области ИИ, включая Оксфорд, Кембридж и Имперский колледж Лондона; расширение доступа к развитию цифровых и STEM-компетенций, в том числе программ Skills for Life и Get the Jump; обучение от компаний, включая Nvidia; программы подготовки управленческих кадров Fast Stream; программы по привлечению талантов из-за рубежа, система визовых возможностей для высококвалифицированных кадров (High-Skilled Visa System)

Анализ стратегий США, Китая и Великобритании показывает, что все три страны рассматривают ИИ как ключевой элемент глобального лидерства, но реализуют его через разные институциональные и финансовые модели. США выстраивают стратегию на основе рыночного стимулирования, сочетая масштабные государственные инвестиции, дерегуляцию и приоритет национальной безопасности. Китай опирается на государственное планирование и мобилизацию частного сектора, совмещая директивное управление, инвестиции в инфраструктуру и массовую подготовку кадров, что отражает модель централизованной технологической мобилизации. Великобритания делает ставку на академическую базу и гибкое регулирование, формируя «проинновационный» подход, который должен компенсировать ограниченный уровень инвестиций по сравнению с США и Китаем.

В кадровой политике также проявляются различия. США концентрируются на масштабных образовательных программах – от школьного уровня до университетских грантов и профессиональной подготовки в военной и гражданской сферах. Китай развивает систему университетов, академий и грантовых инициатив, сочетая внутреннюю подготовку и привлечение зарубежных специалистов. Великобритания акцентирует внимание на сохранении и привлечении талантов, использовании своей академической репутации, а также на создании гибких визовых режимов для специалистов.

Таким образом, США демонстрируют модель геополитической мобилизации через контроль над инфраструктурой и стандартами, Китай – модель технологической самодостаточности и массового внедрения ИИ в экономику и общество, а Великобритания – стратегию «гибкого регулирования» с упором на этику, образование и интеграцию в международное сотрудничество. Эти подходы позволяют странам не только укреплять внутренние экосистемы ИИ, но и формировать различные модели глобального влияния.

Если лидеры глобальной цифровой трансформации демонстрируют устойчивые позиции в международных рейтингах, то страны постсоветского пространства находятся на ином этапе формирования национальных экосистем ИИ. Для них характерны более низкие значения по индексам инновационного развития, цифрового государства и готовности к использованию ИИ. Стратегии развития ИИ в этих странах различаются по приоритетам и глубине проработки, но объединяет их стремление встроиться в глобальный технологический процесс.

В России ключевым ориентиром выступают технологический суверенитет и национальная безопасность. Национальная стратегия до 2030 г. предполагает импортозамещение, развитие собственных мощностей, под-

держку отечественных разработчиков и интеграцию технологий в госуправление и оборонный комплекс. Казахстан связывает развитие ИИ с цифровой трансформацией экономики. Здесь создаются суперкомпьютер, национальная языковая модель KazLLM, международный центр Alem.AI. Особое внимание уделяется подготовке кадров: расширяются университетские и школьные программы, внедряются курсы переподготовки. К 2027 г. планируются многократный рост числа отечественных ИИ-продуктов и их активное внедрение в бизнес и госструктуры.

В Узбекистане ИИ рассматривается как инструмент ускорения экономического роста и повышения качества госуслуг. Стратегия до 2030 г. предусматривает формирование национального центра ИИ, сети лабораторий и внедрение технологий в энергетику, агросектор, здравоохранение и транспорт. Государство стимулирует частные инвестиции и активно развивает образовательные программы. Армения делает ставку на интеграцию в глобальную экосистему и развитие человеческого капитала. Фонд FAST запустил проект Firebird с NVIDIA и Mistral AI, создаются исследовательские центры и образовательные инициативы, что позволяет компенсировать небольшой масштаб рынка через международное сотрудничество.

В целом Россия ориентируется на оборонно-государственную модель и суверенитет, Казахстан и Узбекистан – на цифровизацию экономики, а Армения – на интеграцию и образование. Несмотря на различия, все страны признают стратегическую значимость ИИ для долгосрочного социально-экономического развития.

Направления развития национальной экосистемы ИИ в Беларуси

Формирование национальной экосистемы ИИ в Беларуси требует тесной координации между государством, наукой и бизнесом, что находит свое отражение в выстраивании единой институциональной стратегии. Отправной точкой стало подписание модельного закона «О технологиях искусственного интеллекта», принятого Межпарламентской ассамблеей СНГ, что определило ИИ как стратегический приоритет на государственном уровне. На основе данного нормативно-правового акта будет разработан новый документ для белорусского законодательства – Концепция развития искусственного интеллекта, с целью задавать общие векторы и цели развития [9]. Параллельно с этим в Беларуси действует Декрет Президента Республики Беларусь от 21 декабря 2017 г. № 8 «О развитии цифровой экономики», который, хотя и не посвящен исключительно ИИ, создал беспрецедентно благоприятные условия для всей IT-отрасли, являющейся естественной средой для развития ИИ. Декрет предоставил резидентам ПВТ особые правовые преференции, включая льготный налоговый режим, особый порядок регулирования трудовых отношений и валютных опера-

ций, что напрямую формирует благоприятную финансовую стратегию для компаний, работающих в сфере ИИ.

Институциональная стратегия дополняется активной ролью Национальной академии наук, а именно ОИПИ НАН Беларуси, который консолидирует научно-исследовательский потенциал ведущих университетов, создавая основу для генерации фундаментальных знаний и прорывных исследований. При этом стратегия подразумевает активную работу по разработке этических и правовых норм, регулирующих использование ИИ, что необходимо для построения доверия и интеграции технологий в социально-экономическую жизнь страны.

Неотъемлемой частью этой работы является финансовая стратегия. Основной поток ресурсов традиционно исходит от частного сектора, в первую очередь от крупных резидентов ПВТ, которые инвестируют значительные средства в НИОКР и коммерциализацию продуктов на основе ИИ. Однако для поддержки более рискованных и долгосрочных проектов, особенно в научной среде, критически важную роль играет государственное финансирование через целевые программы и гранты, выделяемые как БРФФИ, так и в рамках государственных заказов на разработку технологий для приоритетных отраслей: здравоохранения, сельского хозяйства и логистики. Главным элементом финансовой стратегии являются также привлечение венчурного капитала и создание специальных фондов, соинвестирующих в стартапы на ранних стадиях, что позволяет преодолеть разрыв между научной идеей и ее рыночной реализацией, обеспечивая финансовый мост для инноваций.

Ключевым звеном является эффективная кадровая стратегия, нацеленная на формирование и удержание конкурентоспособного человеческого капитала. Ее фундамент закладывается в системе высшего образования, где ведущие вузы страны активно открывают новые специализации и магистерские программы по машинному обучению, анализу данных, стремясь удовлетворить растущий спрос со стороны рынка. Однако понимая, что скорость развития ИИ опережает возможности академических программ, экосистема делает серьезную ставку на программы дополнительного образования и профессиональной переподготовки. Главным вызовом стратегии остается проблема «утечки мозгов», для противодействия которой создаются условия для интересной работы и карьерного роста внутри страны, чтобы талантливые специалисты видели Беларусь как привлекательное место для реализации своих амбиций в сфере разработки ИИ.

Заключение

Развитие национальной экосистемы ИИ в Беларуси должно основываться не на прямом заимствовании зарубежных моделей, а на формирова-

нии собственной, адаптированной к национальным условиям стратегии. Мировая практика показывает, что стратегия, эффективная в определенное время и в конкретной среде, не всегда может быть успешно имплементирована в других государствах. Для нашей страны ключевым становится поиск гибридного подхода, сочетающего внутренние институциональные возможности и международное сотрудничество.

Прежде всего следует признать данные новым фактором производства наряду с трудом, капиталом и землей. Создание и поддержка национальных дата-хранилищ, а также формирование правовой базы для регулирования доступа и оборота данных должны стать основой технологического развития. В этом контексте стратегической задачей выступает создание доверительной среды для бизнеса и общества, где использование ИИ опирается на прозрачные правила и этические стандарты. Не менее важна технологическая безопасность. В условиях санкционного давления и ограниченного доступа к критическим технологиям Беларусь должна развивать собственные вычислительные мощности, поддерживать импортонезависимые решения и формировать резервы технологического суверенитета. Это позволит снизить зависимость от внешних акторов и гарантировать устойчивость национальной цифровой инфраструктуры.

С учетом ограниченности ресурсов Беларусь не может соперничать с мировыми лидерами в масштабах инвестиций или объемах кадрового резерва. Однако страна способна сконцентрироваться на нишевых направлениях, например, в разработке отраслевых решений для здравоохранения, сельского хозяйства, логистики и государственного управления. Особую перспективу представляет интеграция в более широкие региональные инициативы. Сотрудничество в рамках Союзного государства может стать аналогом практик Европейского союза: объединение ресурсов, унификация стандартов, совместные исследовательские центры и трансграничные цифровые платформы. Такой формат позволит Беларуси выйти за пределы национального масштаба и интегрироваться в более крупную экосистему, сохраняя при этом собственные стратегические приоритеты.

Таким образом, для нашей страны оптимальным решением станет не копирование чужих моделей, а формирование адаптивной стратегии, основанной на обеспечении технологической безопасности, развитии нишевых решений и углублении сотрудничества в рамках региональной интеграции. Данный подход позволит стране укрепить свои позиции в глобальной конкуренции, избежав рисков догоняющей модели.

Список использованных источников

1. Влияние технологий искусственного интеллекта на экономику и бизнес [сайт] // TAdviser. – URL: https://www.tadviser.ru/index.php/ (дата обращения: 01.10.2025).

2. Strategic Approach to Artificial Intelligence. – URL: https://www.english.digmin.dk/Media/638719220318136690/Stategic%20Approach%20to%20Artificial%20Intelligence.pdf (date of access: 01.10.2025).

3. National AI Strategy // OECD. – URL: https://oecd.ai/en/dashboards/policy-initiatives/national-ai-strategy-20-4102 (date of access: 01.10.2025).

4. National AI Strategy // Smart Nation Singapore. – URL: https://www.smartnation.gov.sg/initiatives/national-ai-strategy (date of access: 01.10.2025).

5. U.S. AI.gov. – URL: https://www.ai.gov/ (date of access: 01.10.2025).

6. Национальные стратегии развития искусственного интеллекта до 2030 года [сайт] // IT-World. – URL: https://www.it-world.ru/it-news/3a5im0am3auc8wsoc0kg8kwockskw4k.html (дата обращения: 01.10.2025).

7. Artificial Intelligence in China // Halyk Finance. – URL: https://halykfinance.kz/download/files/analytics/AC_ai_china.pdf (date of access: 01.10.2025).

8. How the UK Could Lead the World in AI // EU Reporter. – URL: https://ru.eureporter.co/internet-2/artificial-intelligence/2025/03/15/how-the-uk-could-lead-the-world-in-ai (date of access: 01.10.2025).

9. В Беларуси планируют разработать концепцию развития искусственного интеллекта [сайт] // БЕЛТА. – URL: https://belta.by/society/view/v-belarusi-planirujut-razrabotat-kontseptsiju-razvitija-iskusstvennogo-intellekta-731272-2025/ (дата обращения: 01.10.2025).

УДК 004.89

Классификация займа с помощью методов машинного обучения

В. И. Бегунков✉, **М. Я. Ковалев**
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: vbegunkov@gmail.com

Введение

Кредитование давно стало важным инструментом увеличения скорости роста экономик как развитых, так и развивающихся стран. В связи с тем что в любой экономике имеются субъекты хозяйствования с избытком и недостатком денежных средств, кредитование позволяет распределять денежные активы с максимальной выгодой для всех участников, а значит, максимально использовать потенциал экономики отдельно взятой страны. Ключевая роль в этом процессе принадлежит финансовым институтам, которые должны эффективно распределять денежные активы между субъектами в рамках предоставления финансовых услуг. Поэтому финансовым организациям как никогда важно иметь возможность наиболее точно определять надежных заемщиков. При этом потенциальные клиенты ожидают инновационных и индивидуальных решений в данной сфере, которые должны уменьшить время принятия решений по выдаче займов [1]. Учитывая вышесказанное, финансовые институты нуждаются в современных инструментах, которые позволяют удаленно и максимально быстро принимать решения по поступающим запросам на предоставление займов. Машинное обучение является одним из вариантов, на основе которого такие инструменты принятия решений могут быть разработаны.

Описание данных

В качестве данных для настройки параметров и проведения экспериментов с рассматриваемыми методами искусственного интеллекта используются исторические данные по выданным на платформе для кредитования от человека человеку Lending Club займам, состоящие из 2 260 668 строк (займы, выданные за период с апреля 2016 по сентябрь 2018 г.). Перечень входных показателей и принцип преобразования входных данных описаны в работе [2]. Таким образом, финальный набор входных данных состоит из m = 1 221 731 позиций и n = 54 входных показателей.

Оценка эффективности методов осуществляется на основе коэффициента эффективности *A (Accuracy)*, точности *P* (*Precision*), полноты *R* (*Recall*) и меры F_1 [3].

Классификация займа с использованием логистической регрессии

При использовании алгоритма логистической регрессии для решения задачи классификации займа [2] выявлено, что применение нормализации улучшает точность прогнозирования. Также обнаружено, что абсолютная сбалансированность целевых значений не привела к улучшению конечных результатов. Установлено, что оптимальным граничным значением для алгоритма логистической регрессии является 0,49 вместо используемого по умолчанию 0,5. Определено, что увеличение показателей в наборе входных данных и их преобразование с помощью средних и медианных значений приводят к улучшению точности прогнозирования, однако увеличение количества входных показателей с использованием полиномов не привело к однозначному улучшению показателей модели, но существенно увеличило длительность обучения задействованного алгоритма. Применение метода главных компонент позволило получить максимальное значение коэффициента эффективности *A* и является целесообразным при решении задачи классификации займа при использовании логистической регрессии.

Классификация займа с использованием нейронной сети прямого распространения

При исследовании применения нейронной сети прямого распространения для решения задачи классификации займа [4] обнаружено, что использование нормализации тоже улучшает точность прогнозирования. Также выявлено, что поиск оптимального числа нейронов в скрытом уровне и оптимального значения α (размер шага градиентного спуска) привел к улучшению стоимостной функции и коэффициента эффективности *A*. Однако абсолютная сбалансированность целевых значений не привела к улучшению конечных результатов выбранных метрик. При этом установлено, что оптимальным граничным значением для выходного уровня данной нейронной сети является 0,51 вместо используемого по умолчанию 0,5. Было определено, что применение функции активации *ReLu* привело к улучшению результатов классификации. Обнаружено также, что увеличение показателей в наборе входных данных и их преобразование с помощью средних, модальных и медианных значений послужили улучшению стоимостной функции и коэффициента эффективности *A*. В то же время применение полиномов и метода главных компонент не привело к увеличению коэффициента эффективности *A* на тестовых данных. Опти-

мальные значения коэффициента эффективности *A*, меры F_1 и метрики *P* оказались выше значений, рассчитанных с помощью логистической регрессии.

Классификация займа с использованием глубокой нейронной сети прямого распространения

В данном исследовании рассмотрено использование глубокой нейронной сети прямого распространения для решения задачи классификации займа [5]. В результате определена лучшая структура изучаемой нейронной сети. Однако при этом определено, что альтернативное использование инициализации «*He*» не привело к улучшению выбранной метрики модели. Также установлено, что использование мини-пакетного градиентного спуска не привело к улучшению значения величины *A*. В свою очередь, при задействовании альтернативных методов оптимизации глубокой нейронной сети было выявлено, что использование градиентного спуска с импульсом привело к улучшению модели. Вместе с тем обнаружено, что осуществление поиска альтернативного разбиения данных на тренировочный и тестовый наборы является целесообразным, так как это привело к улучшению модели на основе глубокой нейронной сети по выбранной метрике. Также определено, что использование метода исключения на нулевом уровне глубокой нейронной сети привело к улучшению коэффициента эффективности *A*. Итоговая точность прогнозирования при использовании глубокой сети прямого распространения оказалась выше полученной при использовании нейронной сети прямого распространения и логистической регрессии, рассмотренных ранее.

Классификация займа с использованием алгоритма случайного леса

Во время изучения использования алгоритма случайного леса для решения задачи классификации займа установлено [6], что поиск лучшего максимального количества входных показателей привел к улучшению прогнозирования модели. Также выявлено, что от количества деревьев в алгоритме случайного леса зависит точность прогнозирования и целесообразно проводить поиск лучшего количества деревьев в ансамбле при решении конкретной задачи. Это в итоге привело к улучшению прогнозируемых результатов. В завершение поиск альтернативного разделения исходных данных на тренировочный и тестовый наборы привел к увеличению коэффициента эффективности *A* на тестовых данных.

Сравнительный анализ использования логистической регрессии, нейронной сети прямого распространения, глубокой нейронной сети прямого распространения и алгоритма случайного леса для решения задачи классификации займа

Как следует из результатов исследования [6], при использовании алгоритма случайного леса наибольшее значение коэффициента эффективности A на тестовых данных было 80,55 %. Целесообразно сравнить полученные результаты со значениями, рассчитанными при использовании логистической регрессии [2], нейронной сети прямого распространения [4] и глубокой нейронной сети прямого распространения [5]. Из табл. 1 видно, что с помощью глубокой нейронной сети прямого распространения получено наивысшее (лучшее) значение коэффициента эффективности A. При этом важно подчеркнуть, что средняя длительность обучения глубокой нейронной сети прямого распространения существенно выше аналогичных величин при использовании других алгоритмов, указанных в табл. 1.

Таблица 1

Сравнение лучших результатов при применении логистической регрессии, нейронной сети прямого распространения, глубокой нейронной сети прямого распространения и алгоритма случайного леса

Результаты	При алгоритме случайного леса	При глубокой нейронной сети прямого распространения	При нейронной сети прямого распространения	При логистической регрессии
Средняя длительность обучения, с	1520	151 681	17 191	2063
Среднее значение стоимостной функции	не применимо	0,442 56	0,443 48	0,457 06
Accuracy training, %	100	80,40	80,43	79,93
Accuracy testing, %	80,56	80,65	80,32	80,04

Значения метрик F_1, P и R при лучшем коэффициенте эффективности A сравниваемых алгоритмов по невозвратным и возвратным займам, а также при расчете средневзвешенной величины отражены в табл. 2.

Таблица 2

Мера F_1, *P (Precision)* и *R (Recall)* метрики при применении логистической регрессии, нейронной сети прямого распространения, глубокой нейронной сети прямого распространения и алгоритма случайного леса

Класс	Алгоритм случайного леса			Глубокая нейронная сеть прямого распространения			Нейронная сеть прямого распространения			Логистическая регрессия		
	F_1	*P*	*R*	F_1	*P*	*R*	F_1	*P*	*R*	F_1	*P*	*R*
Невозвратные займы, %	15,40	60,49	8,82	18,21	59,79	10,74	20,96	56,45	12,87	15,20	54,78	8,83
Возвратные займы, %	89,02	81,16	98,56	89,03	81,43	98,19	88,76	81,48	97,48	88,69	80,89	98,15
Средневзвешенное, %	74,25	77,02	80,56	74,82	77,09	80,65	75,02	76,41	80,33	73,79	75,60	80,04

Как следует из табл. 2, средневзвешенные значения метрик *P* и *R* оказались выше для глубокой нейронной сети прямого распространения при использовании в качестве критерия оптимизации коэффициента эффективности *A*. При этом, как подчеркивалось выше, результат может быть иным, если в качестве критерия оптимизации использовать другой критерий.

Заключение

Из результатов сравнительного анализа следует, что итоговая точность прогнозирования при использовании глубокой сети прямого распространения [5] оказалась выше полученной при использовании нейронной сети прямого распространения и логистической регрессии, рассмотренных в предыдущих исследованиях [2, 4], а также выше значений, полученных при использовании алгоритма случайного леса [6]. Однако стоит отметить, что решение задачи классификации займа с использование алгоритма случайного леса требует существенно меньше производительных ресурсов и времени, а точность прогнозирования лишь несколько уступает полученной с помощью глубокой нейронной сети прямого распространения и вы-

ше, чем точность, полученная при использовании как логистической регрессии, так и нейронной сети прямого распространения. Исходя из полученных результатов исследования, сформулируем лучшую стратегию решения задачи классификации займа: вначале следует использовать и исследовать алгоритм случайного леса, так как при существенно меньших требуемых ресурсах (вычислительных мощностей и времени обучения) точность прогнозирования уступает лишь полученной с помощью глубокой нейронной сети. После того как разработан и функционирует инструмент классификации займа на основе алгоритма случайного леса, целесообразно продолжить исследования классификации займа с использованием глубокой нейронной сети прямого распространения, что в среднесрочной перспективе позволит улучшить точность прогнозирования.

Список использованных источников

1. Gerhard, F. The coming opportunity in consumer lending / F. Gerhard, A. Harlalka, R. Suvanam // McKinsey Quarterly. – 2021. – URL: https://www.mckinsey.com/business-functions/risk-and-resilience/our-insights/the-coming-opportunity-in-consumer-lending (date of access: 01.05.2021).

2. Бегунков, В. И. Классификация займов с использованием логистической регрессии / В. И. Бегунков, М. Я. Ковалев // Информатика. – 2023. − Т. 20, № 1. – С. 55–74. – DOI: 10.37661/1816-0301-2023-20-1-55-74.

3. Metz, C. E. Basic principles of ROC analysis / C. E. Metz // Seminars in Nuclear Medicine. – 1978. – Vol. 8, no. 4. – P. 283–298.

4. Бегунков, В. И. Классификация займа с использованием нейронной сети прямого распространения / В. И. Бегунков // Информатика. – 2024. − Т. 21, № 1. – С. 83–104. – DOI: 10.37661/1816-0301-2024-21-1-83-104.

5. Бегунков, В. И. Классификация займа с использованием глубокой нейронной сети прямого распространения / В. И. Бегунков, М. Я. Ковалев // Весці Нацыянальнай акадэміі навук Беларусі. Серыя фізіка-матэматычных навук. – 2025. – Т. 61, № 2. – С. 159–174. – DOI: 10.29235/1561-2430-2025-61-2-159-174.

6. Бегунков, В. И. Классификация займа с использованием алгоритма случайного леса и сравнительный анализ с другими классификаторами / В. И. Бегунков, М. Я. Ковалев // Доклады Национальной академии наук Беларуси. – 2025. – Т. 69, № 2. – С. 101–108. – DOI: 10.29235/1561-8323-2025-69-2-101-108.

УДК 005

Квантовая загадка искусственного интеллекта

Н. П. Беляцкий
Белорусский государственный
экономический университет, Минск

В технологиях XXI в. перемен больше, чем за предыдущие столетия бурного развития. Новые технологии, научные исследования, теории возникновения и развития жизни на планете изменили людей вместе со скоростями существования и восприятия как самих себя, так и других. Три закона Исаака Ньютона, которые в свое время дали мощный толчок для механистического изучения природы явлений, где все определенно и предсказуемо, уже не дают объяснений и миропонимания.

Новая квантовая (цифровая, электронная) эпоха базируется на неопределенном, относительном, неизвестном и вероятностном, когда неожиданное случается не реже ожидаемого. Непрерывное последовательное развитие характерно для старого менеджмента. Цифровая эпоха включает (наряду с новым представлением всей картины организационного развития) квантовые дискретные скачки, кризисы, неожиданные прорывы в новые элементные базы технологического развития. Квантовое представление управленческих отношений, когда одни люди управляют другими, когда субъекты и объекты управления меняются местами, включает неявно связанные, запутанные системы, когда взаимосвязи возникают и спонтанно, и закономерно. Закономерности возникают в процессе взаимодействия элементов управленческой системы. Как отмечают физики, организационная простота классической механики «уступает дорогу самоорганизующейся сложности» квантовой эпохи.

Наиболее характерным показателем этой эпохи стал интернет. Но развитие квантовых технологий (чипы, лазеры, смартфоны, планшеты, роботы, компьютеры) еще более убедительно характеризуют данную эпоху. Не менее острой ее характеристикой стал компьютеризированный загадочно интегрированный интеллект на основе синтеза знаний благодаря квантовым технологиям работы с большими данными. Такое новое порождение естественного интеллекта получило название искусственный интеллект (ИИ).

Ученые утверждают, что наш мир на всех уровнях имеет квантовую природу (Ведрал, В. Жизнь в квантовом мире / В. Ведрал; пер. с англ.

А. А. Сорокина // В мире науки. – 2011. – № 8). Такое восприятие картины мира, его природа прежде всего подразумевают подход к явлениям с позиции «и-и» вместо «или-или», т. е. многообразие возможных решений. Квантовое понимание управленческих процессов, как и всех остальных, включая ИИ, акцентирует внимание на взаимодействии, целостном единстве, вопросах, а не на ответах, на интегрированном подходе, потенциальных возможностях и вероятностях, лидерстве, а не на руководстве.

Согласно квантовой физике результат исследования зависит от выбора инструмента: поток электронов или последовательность волн. Иначе говоря, ответ зависит от самого вопроса. Есть и то, и другое, но не одновременно. В этом суть принципа неопределенности. Применительно, например, к бизнесу: смысл и в создании рабочих мест, и в получении прибыли. Есть и то, и другое, но если тиражируешь бизнес, то приоритетны новые места и только потом прибыль.

Потенциальный квантовый лидер обязан найти свое предназначение и понимание того, что для него приоритетно. Проблема расстановки приоритетов (а значит, и смыслов) не столько умственная, сколько эмоциональная и даже духовная. Роль ИИ здесь противоречива и ограничена.

С позиции классической физики И. Ньютона законы природы не включают умственные, эмоциональные и психологические аспекты. Последователи ньютоновского мышления в лице Фрейда и Смита опирались при исследованиях поведения человека и экономики на законы природы и классической физики. Британец Ф. Тейлор, француз А. Файоль, как утверждают современные ученые, также полагали, что любая организация и ее участники, т. е. люди, подчиняются определенным фундаментальным законам и принципам. Стоит их познать, и можно контролировать как динамику рынка, так и поведение людей, иначе говоря, можно программировать экономические и управленческие процессы, а также контролировать и планировать. Конечно, это оказалось верным в той степени, в которой сотрудники и организации обладают качествами механизма.

Квантовые физики так отмечают данную особенность: «Мир бизнеса больше не нуждается в человеческих “мыслящих машинах”. Кремниевые микросхемы дешевле, быстрее и надежнее. В свою очередь, “интеллектуальный капитал”, который стремится развивать современные компании, обладает теми человеческими характеристиками, какие не способна воспроизвести машина, теми особыми качествами человека, что не подлежат программированию» [1].

Компьютеры не имеют ни эмоционального, ни духовного интеллекта. У них нет желаний и стремлений к успеху, нет искусственного творчества. Программы преобразований не могут обеспечить развитие этих качеств. Между тем нет таких программ и для развития человеческих ресурсов,

эмоционального и духовного интеллекта, креативности и интуиции. Но даже при их наличии у креативных консультантов организационного развития доминирует механический, условно говоря «атомарный» подход на уровне сокращения персонала и тех механизмов изменений, которые не затрагивают изменения среды для благоприятного развития эмоционального интеллекта и соответствующего стиля лидерства.

Иначе говоря, механистический подход заложен в самом понятии «механизма изменений», который не затрагивает формирование потенциала преобразований существующей культуры и форм управленческих отношений. Например, классическая механика, основанная на законах И. Ньютона, не затрагивает процессы внутри атомов, никак не рассматривает ядерные силы и силы взаимодействия элементарных частиц, из которых состоят эти атомы.

Механистические взгляды на концепции руководителей, предпринимателей, бизнес-лидеров не позволяют оценить их личностные особенности по соответствующим критериям как-будто есть такие поддающиеся измерению критерии. Проблема в том, что все личностные особенности, все уровни потребностей (например, пирамида А. Маслоу) взаимопроникают, а потребности неотделимы, и только условно их можно ранжировать. Все наши компетенции поддерживают друг друга в динамике, их приоритеты меняются. Непонимание целостности и системного аспекта организационного развития (кибернетических основ) есть непонимание квантовых предпосылок или основ квантовой динамики как ИИ, так и организации. Например, потребности смысла, ценностные ориентации, миссии (духовные потребности) наполняют содержанием всех остальных (экономические, материальные, потребности саморазвития), но не ИИ.

Исходным уровнем креативного менеджмента можно считать миссию, ценностные ориентации, базовые видения организации и персонала. Это квантовый уровень творческого мышления, на котором мышление человека напоминает квантовые процессы ИИ и наоборот. Это эмоционально-интуитивный духовный уровень креативности. Его можно назвать квантовым уровнем творческого мышления. Здесь аналогия менеджмента с квантовой физикой в том, что данная физика объясняет, исследует уровень реальности в ситуациях неопределенности, неясности, вероятности поведения элементарных частиц энергии и вещества, причем не каждой в отдельности, а только в составе с другими. Аналоги энергии и вещества на уровне личности – это его материальное воплощение. Они взаимопроникают умственной, эмоциональной и духовной составляющими.

Неопределенность, вероятность квантового менеджмента и детерминированность, устойчивость классического тейлоровского, файолевского менеджмента находятся (балансируются) на грани. В теории хаоса гранью

называется критическое состояние системы, при котором ее минимальное изменение вызывает либо неуправляемое (хаотическое) поведение системы, либо замкнутость системы в данной критической точке. Ученые определяют грань хаоса как границу между порядком и хаосом, стабильностью и нестабильностью. Здесь возникает самоорганизация, чтобы выжить. Так на грани балансируют природные системы.

Хаотические системы в природе (непредсказуемые погодные изменения, водовороты речных течений, перепады кровяного давления в организме, настроение и т. д.) балансируют, поддерживают равновесие (что особенно важно в самолетостроении и воздушном плавании) между устойчивостью и неустойчивостью, предсказуемостью и наоборот. Природные системы всегда готовы к переменам, т. е. это открытые, динамические, самоорганизующиеся системы. Конструкторы роботов, разработчики ИИ изучают строение простейших животных, а также их поведение, чтобы потом заимствовать полученные результаты в своей работе над «электронно-механическими системами». Точно также можно получить успех, проводя аналогию между квантовой механикой и классическим менеджментом.

Приведем пример такой аналогии из опыта реорганизации корпорации «Волво». До этого группы инженеров были жестко структурированы, постоянны и централизованы. Каждый член группы отвечал за определенную операцию и строго придерживался отведенной ему роли, чем обеспечивалась регулярная, детерминированная структура всего коллектива. Для обеспечения выпуска желаемого продукта группы организовывались по иерархическому принципу. Но потом компания, вооружившись новой идеей, все начала с нуля. Инженеры работают в группах без жесткой организации, в которые входят также сотрудники других специальностей. Подгруппы входят в состав более крупных коллективов. Инженеры могут свободно переходить из одной группы или подгруппы в другую. На рабочих местах не только стулья, но и столы были оснащены колесиками [1].

Квантовые физики учат нас работать с учетом действия принципов неопределенности, дискретности и неоднозначности.

Теории нечеткой логики эволюционных алгоритмов, когнитивных карт, нечетких игр, нечетких множеств, когнитивных сетей, сетей доверия – только один их перечень заставляет задуматься о значении квантовых предпосылок для принятия управленческих решений, а значит, и для менеджмента. Проблемы менеджмента находятся в нехватке необходимой информации, которую надо искать в переизбыточных объемах данных. Информационно-компьютерные и интеллектуально-аналитические технологии ИИ успешно обрабатывают большие массивы данных, в том числе низкоструктурированных. Однако в цифровом бизнесе стоит проблема скорости получения именно релевантной информации. Эту проблему ре-

шает ИИ путем когнитивного анализа нечетких, расплывчатых, неопределенных данных. Когнитивная аналитика включает как интеллектуальный поиск, так и экспертную поддержку принятия управленческих решений. Далее напрашивается квантовая аналитика такой поддержки на основе эмоционального интеллекта и параллельного мышления. Путь от когнитивного менеджмента к квантовому менеджменту и есть развитие ИИ на основе неопределенностей, вероятностей, возможностей и потребностей.

Интеллектуальный анализ данных включает «раскопку данных» (Data Mining) на основе обнаружения релевантной информации в «скрытых» данных для принятия решений. Такой анализ использует квантовые методы поиска ранее неизвестных, нетривиальных источников данных по формуле «и то – и то». На пути от данных к информации, а затем к знаниям, известны следующие шаги: формирование классов данных – установление взаимосвязи между ними – ассоциативное, квантовое мышление для поиска зависимости между случайными событиями – поиск причин отклонений от ранее принятых решений. Интеллектуальный анализ данных получил название Knowledge Discovery in Databases, KDD[1].

С разработками квантовых компьютеров, дополнением к ним и логическому мышлению эмоционального параллельного мышления становятся возможным суперскоростной квантовый анализ больших данных и на его основе принятие управленческих решений благодаря ИИ. Данное направление развития креативного мышления трансформировалось в квантовый менеджмент цифрового бизнеса успешных компаний.

Осмысление тенденции развития менеджмента, как представляется ученым квантового направления физики, требует понимания универсальности принципов неопределенности, дискретности и виртуальности. Такой подход предлагает, в частности, предпринимательскую креативность по формуле: возможность (приоритетны анализ, логика, последовательное мышление) – потребность (приоритетны ассоциативность, интуиция, параллельное мышление) – необходимость (квантовый мотив, духовный интеллект).

Квантовая организация приоритетно ориентирована на взаимодействие с внешней средой, объединение маркетинговых ресурсов с другими компаниями для совместных усилий по развитию бизнеса, например путем создания расширяющейся базы данных. Квантовая организация ориентирована на интеграцию и сотрудничество. Ей не свойственны идеалы контроля и конкуренции. Это «коллектив коллективов», «помощь для самопомощи», «рыба в воде, а вода в рыбе». Такая организация находится в рав-

[1]Цифровой бизнес : учебник / под ред. О. В. Китовой. – М. : ИНФА-М, 2019. – С. 281.

новесии с внешней средой, представители которой, включая конкурентов, находятся в процессе развития нематериальных активов и, в частности, интеллектуального капитала наряду с персоналом организации, а также поставщиками и потребителями. В основе такой довольно странной и противоречивой практической концепции лежат неопределенность и неоднозначность, когда прогнозирование и контроль не достигают своего назначения. Принципы действия бизнеса вытекают из его внешней и внутренней среды, а она отличается высокой динамикой, неопределенностью, часто состоянием на грани хаоса. Чтобы извлекать пользу из работы в такой среде, нужно меняться быстрее, чем она сама, т. е. действовать на опережение. Здесь знания не тождественны контролю. Наличие всех необходимых данных не гарантирует безошибочность математических расчетов развития экономических процессов в отличие от физического процесса на атомарном уровне, т. е. «на уровне законов ньютоновской науки», по действующую принципу детерминизма и причинно-следственных связей. Топ-менеджеры классической компании стремятся избежать непредвиденных событий, чтобы управлять обстоятельствами и даже «управлять непредвиденным». Для этого разработаны концепции опережающего управления, когда ищут ответы на вопросы, которые еще не возникли. Для этого разработаны различные виды, которые заимствованы из технической механики.

Но вот что интересно, чем сильнее контроль, тем меньше его значение для развития системы. На этой счет в кибернетике есть теорема фон Ферстера. Она гласит: «чем жестче связь между элементами системы, тем меньше каждый из них влияет на систему в целом и тем более он отдален от нее. (Кибернетика изучает общие законы мира живой природы и мира машин. Это всеобщая организационная наука, которая определяет принципы организационного развития системы независимо от особенностей ее элементов.) Вывод для практики бизнеса с позиции кибернетических принципов состоит в том, чтобы сделать упор на самоорганизацию всей компании, а не на контроль отдельных функциональных органов, причем эта неопределенность становится взаимной предпосылкой и для ИИ в бизнесе.

Список использованных источников

1. Зохар, Дана. Квантовый лидер: революция в мышлении и практике бизнеса / Дана Зохар; пер. с англ. – М. : ООО Книжное изд. «София», 2017. – 352 с.

2. Уорнер, М. Виртуальные организации. Новые формы ведения бизнеса в XXI веке / М. Уорнер, М. Витцель; пер. с англ. – М. : Добрая книга, 2005. – 296 с.

3. Беляцкий, Н. П. Дискретность лидерства / Н. П. Беляцкий // Менеджмент в России и за рубежом : все о теории и практике управления бизнесом, финансами, кадрами. – 2004. – № 6. – С. 125–130.

4. Беляцкий, Н. П. Квантовая природа менеджмента / Н. П. Беляцкий // Менеджмент в России и за рубежом. – 2005. – № 4. – С. 3–7.

5. Беляцкий, Н. П. Квантовые скачки бизнеса / Н.П. Беляцкий // Наука и инновации. – 2019. – № 11. – С. 64–68.

6. Беляцкий, Н. П. Креативный менеджмент : учебник / Н. П. Беляцкий. – Минск : Вышэйшая школа, 2022. – 384 с.

7. Беляцкий, Н. П. Квантовая философия менеджмента / Н. П. Беляцкий // Беларуская думка. – 2022. – № 11, С. 72–80.

8. Беляцкий, Н. П. Управление персоналом : учебник / Н. П. Беляцкий. – Минск : Вышэйшая школа, 2023. – 463 с.

УДК 681.5.04

Технологически независимая оптимизация функций многозначной логики

П. Н. Бибило✉, **В. И. Романов**
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: bibilo@newman.bas-net.by

Введение

Схемная реализация цифровых систем, проектные спецификации которых представлены на высокоуровневых языках проектирования, традиционно выполняется в два этапа. На первом этапе – высокоуровневом синтезе – осуществляется переход к функционально-структурным логическим описаниям, на втором этапе – логическом синтезе – проектируются (синтезируются) логические схемы. Логический синтез (технологическое отображение) в заданную библиотеку логических элементов СБИС предваряется технологически независимой оптимизацией [1], которая определяет основные параметры логических схем – площадь (число транзисторов), быстродействие, энергопотребление. Система *FLC-2* предназначена для технологически независимой логической оптимизации функционально-структурных описаний цифровых блоков СБИС [2]. В *FLC-2* реализована интеллектуальная поддержка проектирования, основанная на применении продукционно-фреймовой модели представления знаний [3] и соответствующих методах логического вывода.

В данной работе сообщается о технологически независимой оптимизации MDD-представлений функций многозначной логики (*Multi-valued Decision Diagram* – многозначная диаграмма решений), что позволяет расширить комбинированные методы логической оптимизации системы *FLC-2*.

Определения

Пусть переменные x_i, $i = 1, \ldots, n$, принимают значения из множеств $S_i = \{s_i^1, \ldots, s_i^{q_i}\}$ соответственно, а функция f принимает значения из множества $S_f = \{s_f^1, \ldots, s_f^k\}$. Тогда *функция многозначной логики* (*многозначная функция) f* определяется как отображение $f : S_1 \times S_2 \times \ldots \times S_n \to S_f$. *Конечный пре-*

дикат p определяется как отображение $p: S_1 \times S_2 \times ... \times S_n \to \{0, 1\}$. *Частичная (не полностью определенная) многозначная функция* f – это отображение $f: S_1 \times S_2 \times ... \times S_n \to \{S_f \cup \{-\}\}$, а *частичный (не полностью определенный) конечный предикат* p – отображение $p: S_1 \times S_2 \times ... \times S_n \to \{0, 1, -\}$. В этих определениях $S_1 \times S_2 \times ... \times S_n$ – декартово произведение множеств S_1,..., S_n (множество $M = S_1 \times S_2 \times ... \times S_n$ будем называть *пространством*, на котором определена многозначная функция f), а через «-» обозначено неопределенное значение многозначной функции либо конечного предиката. Неопределенное значение многозначной функции может быть заменено элементом из множества $S_f = \{s_f^1, ..., s_f^k\}$, а неопределенное значение конечного предиката – элементом из множества {0, 1}. В зарубежной литературе многозначные функции $f(x_1, x_2, ..., x_n) = f(\boldsymbol{x})$, $\boldsymbol{x} = (x_1, x_2, ..., x_n)$, называются *multi-valued input multi-valued output functions* (англ.). Далее для краткости будем называть переменные (аргументы) таких функций и сами функции *многозначными* (можно было бы сказать также «конечнозначными»), подразумевая, что области значений переменных и функций содержат конечное (больше двух) число значений, и отличая их от булевых (двухзначных) аргументов и функций. Если $S_1 = S_2 = ... = S_n = S_f = \{0, 1\}$, то конечные предикаты являются *булевыми функциями*.

Пример задания функций двухразрядного умножителя

Три различные формы задания функций устройства Mult_2_2 умножения двух целых неотрицательных чисел, заданных двумя двоичными разрядами, приведены в табл. 1. Форма 1 (левая часть табл. 1) – многозначная функция $y = a \times b = ab$, областью значений аргументов *a, b* является множество {0, 1, 2, 3}, функция *y* принимает значения из множества {0, 1, 2, 3, 4, 6, 9}. Форма 2 (средняя часть табл. 1) – система конечных предикатов $y_1 = y_1(a,b)$, $y_2 = y_2(a,b)$, $y_3 = y_3(a,b)$, $y_4 = y_4(a,b)$ отличается от задания многозначной функции тем, что значения функции (числа) закодированы 4-разрядными двоичными кодами.

Форма 3 (правая часть табл. 1) – система булевых функций отличается от системы конечных предикатов тем, что численные значения многозначных аргументов *a*, *b* закодированы 2-разрядными двоичными кодами (a_1, a_2), (b_1, b_2) соответственно. Переход от систем булевых функций, часто считающихся *исходными*, к конечным предикатам и многозначным функциям прост и заключается в кодировании двоичных наборов представляющими их десятичными неотрицательными целыми числами. При кодиро-

вании наборов значений аргументов получаются системы предикатов, последующее кодирование наборов значений предикатов приводит к таблицам задания многозначных функций.

Таблица 1

Различные формы задания функций умножителя Mult_2_2

Многозначная функция			Система конечных предикатов			Система булевых функций		
a	b	y	a	b	$y_1\ y_2\ y_3\ y_4$	$a_1\ a_2$	$b_1\ b_2$	$y_1\ y_2\ y_3\ y_4$
0	0	0	0	0	0 0 0 0	0 0	0 0	0 0 0 0
0	1	0	0	1	0 0 0 0	0 0	0 1	0 0 0 0
0	2	0	0	2	0 0 0 0	0 0	1 0	0 0 0 0
0	3	0	0	3	0 0 0 0	0 0	1 1	0 0 0 0
1	0	0	1	0	0 0 0 0	0 1	0 0	0 0 0 0
1	1	1	1	1	0 0 0 1	0 1	0 1	0 0 0 1
1	2	2	1	2	0 0 1 0	0 1	1 0	0 0 1 0
1	3	3	1	3	0 0 1 1	0 1	1 1	0 0 1 1
2	0	0	2	0	0 0 0 0	1 0	0 0	0 0 0 0
2	1	2	2	1	0 0 1 0	1 0	0 1	0 0 1 0
2	2	4	2	2	0 4 0 0	1 0	1 0	0 4 0 0
2	3	6	2	3	0 1 1 0	1 0	1 1	0 1 1 0
3	0	0	3	0	0 0 0 0	1 1	0 0	0 0 0 0
3	1	3	3	1	0 0 1 1	1 1	0 1	0 0 1 1
3	2	6	3	2	0 1 1 0	1 1	1 0	0 1 1 0
3	3	9	3	3	1 0 0 1	1 1	1 1	1 0 0 1

Бинарные диаграммы решений (BDD-представления)

Система булевых функций из правой части табл. 1 по перестановке $<a_1, a_2, b_2, b_1>$ имеет BDD-представление (рис. 1), которому соответствует следующая совокупность взаимосвязанных формул разложения Шеннона

$$\begin{gathered} y_3 = \overline{a}_1 s_2 \vee a_1 s_3;\ y_1 = a_1 s_0;\ y_2 = a_1 s_1;\ y_4 = a_2 b_2; \\ s_1 = \overline{a}_2 b_1 \vee a_2 s_8;\ s_3 = \overline{a}_2 b_2 \vee a_2 s_{10};\ s_2 = a_2 b_1; \\ s_0 = \overline{a}_2 s_4;\ s_8 = \overline{b}_2 b_1;\ s_{10} = \overline{b}_2 b_1 \vee b_2 \overline{b}_1;\ s_4 = b_2 b_1. \end{gathered} \quad (1)$$

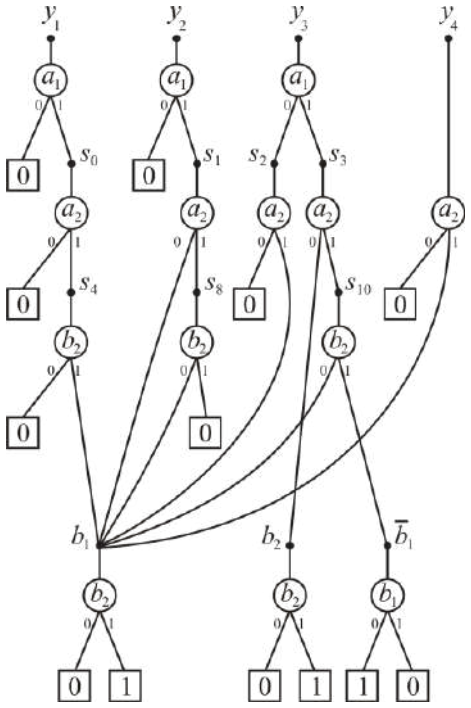


Рис. 1. Бинарная диаграмма решений (BDD)

Суммарное число литералов (булевых переменных либо их инверсий) в правых частях формул (1) равно 30.

Многозначные диаграммы решений и переход к логическим уравнениям

Графовое MDD-представление многозначной функции y из левой части табл. 1 по перестановке <a, b> многозначных переменных показано на рис. 2, графу MDD соответствует совокупность формул (2) разложений Шеннона функции $y = ab$:

$$y = a^{\{0\}}c_0 \vee a^{\{1\}}c_1 \vee a^{\{2\}}c_2 \vee a^{\{3\}}c_3;\ \ c_0{=}0;\ c_1 = b^{\{0\}}0 \vee b^{\{1\}}1 \vee b^{\{2\}}2 \vee b^{\{3\}}3;$$
$$c_2 = b^{\{0\}}0 \vee b^{\{1\}}2 \vee b^{\{2\}}4 \vee b^{\{3\}}6;\ c_3 = b^{\{0\}}0 \vee b^{\{1\}}3 \vee b^{\{2\}}6 \vee b^{\{3\}}9. \qquad (2)$$

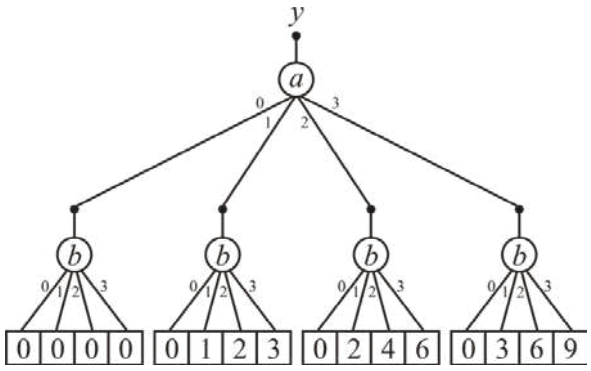


Рис. 2. Многозначная диаграмма решений (MDD)

Запишем в векторном виде уравнения (2), используя кодирование (табл. 2) булевыми переменными различных объектов, входящих в формулы (2):

$$(y_1, y_2, y_3, y_4) = \overline{a}_1\overline{a}_2(c_0^1, c_0^2, c_0^3, c_0^4) \vee \overline{a}_1 a_2(c_1^1, c_1^2, c_1^3, c_1^4) \vee$$
$$\vee a_1\overline{a}_2(c_2^1, c_2^2, c_2^3, c_2^4) \vee a_1 a_2(c_3^1, c_3^2, c_3^3, c_3^4);$$
$$(c_0^1, c_0^2, c_0^3, c_0^4) = (0, 0, 0, 0);$$
$$(c_1^1, c_1^2, c_1^3, c_1^4) = \overline{b}_1\overline{b}_2(0,0,0,0) \vee \overline{b}_1 b_2(0,0,0,1) \vee b_1\overline{b}_2(0,0,1,0) \vee b_1 b_2(0,0,1,1);$$
$$(c_2^1, c_2^2, c_2^3, c_2^4) = \overline{b}_1\overline{b}_2(0,0,0,0) \vee \overline{b}_1 b_2(0,0,1,0) \vee b_1\overline{b}_2(0,1,0,0) \vee b_1 b_2(0,1,1,0);$$
$$(c_3^1, c_3^2, c_3^3, c_3^4) = \overline{b}_1\overline{b}_2(0,0,0,0) \vee \overline{b}_1 b_2(0,0,1,1) \vee b_1\overline{b}_2(0,1,1,0) \vee b_1 b_2(1,0,0,1).$$

Таблица 2

Кодирование многозначных объектов булевыми кодами

Многозначная переменная, функция, кофактор, числовая константа	Булев код
y	(y_1, y_2, y_3, y_4)
a	(a_1, a_2)
b	(b_1, b_2)
$c_0 = 0$	(0,0,0,0)
c_i	$(c_i^1, c_i^2, c_i^3, c_i^4)$
Литералы многозначных переменных	
$a^{\{0\}}$	$\overline{a}_1\overline{a}_2$
$a^{\{1\}}$	$\overline{a}_1 a_2$
$a^{\{2\}}$	$a_1\overline{a}_2$
$a^{\{3\}}$	$a_1 a_2$
$b^{\{0\}}$	$\overline{b}_1\overline{b}_2$
$b^{\{1\}}$	$b_1\overline{b}_2$
$b^{\{2\}}$	$\overline{b}_1 b_2$
$b^{\{3\}}$	$b_1 b_2$

Перейдем от векторных 4-разрядных уравнений к булевым уравнениям, т. е. запишем логические уравнения для каждого разряда:

$$y_1 = \overline{a}_1\overline{a}_2 c_0^1 \vee \overline{a}_1 a_2 c_1^1 \vee a_1\overline{a}_2 c_2^1 \vee a_1 a_2 c_3^1;$$
$$y_2 = \overline{a}_1\overline{a}_2 c_0^2 \vee \overline{a}_1 a_2 c_1^2 \vee a_1\overline{a}_2 c_2^2 \vee a_1 a_2 c_3^2; \; y_3 = \overline{a}_1\overline{a}_2 c_0^3 \vee \overline{a}_1 a_2 c_1^3 \vee a_1\overline{a}_2 c_2^3 \vee a_1 a_2 c_3^3;$$
$$y_4 = \overline{a}_1\overline{a}_2 c_0^4 \vee \overline{a}_1 a_2 c_1^4 \vee a_1\overline{a}_2 c_2^4 \vee a_1 a_2 c_3^4;$$

$$c_0^1 = 0;\ c_0^2 = 0;\ c_0^3 = 0;\ c_0^4 = 0;$$
$$c_1^1 = \bar{b}_1\bar{b}_2 0 \vee \bar{b}_1 b_2 0 \vee b_1\bar{b}_2 0 \vee b_1 b_2 0 = 0;\ c_1^2 = 0;$$
$$c_1^3 = b_1\bar{b}_2 1 \vee b_1 b_2 1 = b_1\bar{b}_2 \vee b_1 b_2;\ c_1^4 = \bar{b}_1 b_2 1 \vee b_1 b_2 1 = \bar{b}_1 b_2 \vee b_1 b_2;$$
$$c_2^1 = 0;\ c_2^2 = b_1\bar{b}_2 1 \vee b_1 b_2 1 = b_1\bar{b}_2 \vee b_1 b_2;\ c_2^3 = \bar{b}_1 b_2 1 \vee b_1 b_2 1 = \bar{b}_1 b_2 \vee b_1 b_2;\ c_2^4 = 0;$$
$$c_3^1 = b_1 b_2 1 = b_1 b_2;\ c_3^2 = b_1\bar{b}_2 1 = b_1\bar{b}_2;\ c_3^3 = \bar{b}_1 b_2 1 \vee b_1\bar{b}_2 1 = \bar{b}_1 b_2 \vee b_1\bar{b}_2;$$
$$c_3^4 = \bar{b}_1 b_2 1 \vee b_1 b_2 1 = \bar{b}_1 b_2 \vee b_1 b_2.$$

После подстановки булевых констант 0 и сокращения логических выражений получим следующие уравнения

$$\begin{gathered} y_1 = a_1 a_2 c_3^1;\ y_2 = a_1\bar{a}_2 c_2^2 \vee a_1 a_2 c_3^2;\ y_3 = \bar{a}_1 a_2 c_1^3 \vee a_1\bar{a}_2 c_2^3 \vee a_1 a_2 c_3^3; \\ y_4 = \bar{a}_1 a_2 c_1^4 \vee a_1 a_2 c_3^4;\ c_1^3 = b_1\bar{b}_2 \vee b_1 b_2;\ c_1^4 = \bar{b}_1 b_2 \vee b_1 b_2;\ c_2^2 = b_1\bar{b}_2 \vee b_1 b_2; \\ c_2^3 = \bar{b}_1 b_2 \vee b_1 b_2;\ c_3^1 = b_1 b_2;\ c_3^2 = b_1\bar{b}_2;\ c_3^3 = \bar{b}_1 b_2 \vee b_1\bar{b}_2; \\ c_3^4 = \bar{b}_1 b_2 \vee b_1 b_2. \end{gathered} \tag{3}$$

Учитывая, что в (3) среди c_i^j имеются одинаковые уравнения, введем обозначения

$$c_3^1 = b_1 b_2 = p_1;\ c_1^3 = c_2^2 = b_1\bar{b}_2 \vee b_1 b_2 = p_2;\ c_3^2 = b_1\bar{b}_2 = p_3;$$
$$c_1^4 = c_2^3 = c_3^4 = \bar{b}_1 b_2 \vee b_1 b_2 = p_4;\ c_3^3 = \bar{b}_1 b_2 \vee b_1\bar{b}_2 = p_5.$$

Результирующие логические уравнения при переходе от уравнений (2) будут иметь следующий вид:

$$\begin{gathered} y_1 = a_1 a_2 p_1;\ y_2 = a_1\bar{a}_2 p_2 \vee a_1 a_2 p_3;\ y_3 = \bar{a}_1 a_2 p_2 \vee a_1\bar{a}_2 p_4 \vee a_1 a_2 p_5; \\ y_4 = \bar{a}_1 a_2 p_4 \vee a_1 a_2 p_4;\ p_1 = b_1 b_2;\ p_2 = b_1\bar{b}_2 \vee b_1 b_2; \\ p_3 = b_1\bar{b}_2;\ p_4 = \bar{b}_1 b_2 \vee b_1 b_2;\ p_5 = \bar{b}_1 b_2 \vee b_1\bar{b}_2. \end{gathered} \tag{4}$$

Суммарное число литералов в уравнениях (4) равно 40. Некоторые уравнения в (4) можно упростить, используя приемы минимизации булевых функций в классе ДНФ – склеивание и поглощение элементарных конъюнкций. Заметим, что для уравнений (1) BDD была использована «лучшая» перестановка булевых переменных разложения, а для MDD – заданная (фиксированная) перестановка многозначных переменных.

Программы

Технологически независимая оптимизация систем многозначных и булевых функций на основе разложений Шеннона выполняется в системе *FLC-2* различными оптимизационными процедурами, основу которых составляют две программы – *BDD_Builder* и *MDD_Builder*.

Программа *BDD_Builder* выполняет BDD-минимизацию систем булевых функций, находя наилучшую перестановку переменных разложения, при которой число различных кофакторов (подфункций) разложения Шеннона является наименьшим. Для систем частичных булевых функций используется так называемое «нулевое» доопределение, когда все неопределенные значения «–» компонентных функций системы заменяются нулями.

В работе [1] проведено широкое экспериментальное исследование данной программы, результаты которого показали высокую эффективность использования *BDD_Builder* для синтеза как схем заказных СБИС, так и программируемых логических интегральных схем *FPGA* (*Field-Programmable Gate Array* – программируемая пользователем вентильная матрица).

Программа *MDD_Builder* минимизирует MDD для заданной перестановки многозначных переменных, после чего по минимизированной MDD получает логические уравнения. Другой режим работы *MDD_Builder* – поиск перестановки многозначных переменных, при которой результирующие логические уравнения будут содержать наименьшее число литералов. Суть алгоритма, реализованного в программе *MDD_Builder* для систем полностью определенных многозначных функций (либо систем предикатов) – поиск одинаковых кофакторов на каждом шаге разложения многозначной функции, для систем частичных функций – поиск доопределений кофакторов для минимизации их числа [4].

Заключение

Логическая оптимизация при схемной реализации функций многозначной логики в системе *FLC-2* может выполняться различными маршрутами: традиционным маршрутом, используя переход (кодирование) от многозначных функций к системам булевых функций, и новым маршрутом, используя минимизацию MDD-представлений многозначных функций и последующий переход к многоуровневым представлениям систем булевых функций в виде логических уравнений. Результаты экспериментальных исследований различных алгоритмов схемной реализации функций многозначной логики с использованием программ *BDD_Builder* и *MDD_Builder* приведены работе в [5].

Список использованных источников

1. Бибило, П. Н. Бинарные диаграммы решений в логическом проектировании / П. Н. Бибило. – М. : Ленанд, 2024. – 560 с.

2. Бибило, П. Н. Система логической оптимизации функционально-структурных описаний цифровых устройств на основе продукционно-фреймовой модели представления знаний / П. Н. Бибило, В. И. Романов // Проблемы разработки перспективных микро- и наноэлектронных систем : сб. тр. / под общ. ред. акад. РАН А. Л. Стемпковского. – М. : ИППМ РАН, 2020. – № 4. – С. 9–16.

3. Бибило, П. Н. Логическое проектирование дискретных устройств с использованием продукционно-фреймовой модели представления знаний / П. Н. Бибило, В. И. Романов. – Минск : Беларус. навука, 2011. – 279 с.

4. Бибило, П. Н. Использование диаграмм решений не полностью определенных функций *k*-значной логики при синтезе логических схем / П. Н. Бибило // Информатика. – 2023. – Т. 20, № 2. – С. 39–64.

5. Романов, В. И. Экспериментальное сравнение алгоритмов схемной реализации функций многозначной логики / В. И. Романов, П. Н. Бибило (в наст. сборнике).

УДК 004.42

Автоматизация принятия решений при анализе состояния сложных объектов и ситуаций

П. Ю. Бранцевич
Белорусский государственный университет
информатики и радиоэлектроники, Минск
E-mail: branc@bsuir.edu.by

Введение

В настоящее время усложняются научно-технические решения и производства, разрабатываются интеллектуальные системы, происходят титанические сдвиги в общественных, социальных, групповых процессах и отношениях. Все это требует автоматизации принятия решений по оценке возникающих ситуаций и явлений.

В производственных процессах значительную часть от эксплуатационных затрат составляют расходы, направляемые на обеспечение работоспособности производственного оборудования. Снизить затраты на его эксплуатацию можно путем внедрения современных систем технического обслуживания, базирующихся на использовании технологий мониторинга, оценки состояния, диагностики, прогнозирования развития дефектов, которые с точки зрения их организации и функционирования являются интеллектуальными системами [1].

Состояние производственного оборудования может характеризоваться многими параметрами основных и вторичных процессов, развивающихся при его работе. Для контроля целесообразно выбирать те, которые достаточно хорошо отражают функциональное состояние объектов и не требуют слишком больших затрат на их определение [1, 2]. На основе анализа состояния группы однотипных механизмов при их функционировании на различных режимах в различном техническом состоянии и на протяжении длительного времени могут быть обоснованы и сформулированы диагностические признаки для локализации мест возникновения дефектов и неисправностей. Это создает условия для построения автоматизированных интеллектуальных систем оценки технического состояния и диагностики, значительно облегчающих работу инженерно-технического персонала [3].

Медицинская диагностика как набор правил, методов и решений, которые позволяют прийти к заключению о наличии или вероятности наличия у человека того или иного заболевания, имеет много общего с техни-

ческой диагностикой [4]. Открытие новых и совершенствование уже существующих методов оценки состояния человека – важное направление медицинских исследований. В связи с этим формализация систем принятия решений на основе получаемых данных о состоянии человека является весьма актуальной, причем все более доступной для создания, по крайней мере для многих частных случаев, за счет растущей вычислительной и информационной мощности применяемых технических средств.

Одним из примеров задачи принятия решений в организационно-управленческой сфере является подбор кадрового резерва руководителей в учреждениях высшего образования, где в качестве параметров используются: возраст, пол, семейное положение, место рождения, образование, вторая ступень высшего образования, наличие ученой степени и ученого звания, стаж работы в занимаемой должности [5].

Модель живого организма как представителя естественного интеллекта, принимающего решения

На сегодняшний день возникает необходимость осуществлять мониторинг состояния не только технических объектов, но и живых организмов, в том числе и человека. Предположим гипотезу, что живой организм – это материальная органическая сущность, состоящая:

– из набора разнообразных первичных преобразователей информации (зрение, слух, обоняние, вкус, рецепторы оценки состояния окружающей среды и собственных органов, нервные окончания и т. д.), преобразующих состояние окружающей среды и самого организма в электрические сигналы, передаваемые в нейронную сеть головного мозга человека;

– органов обеспечения жизнедеятельности, выживания, питания, реакций, работоспособности, воспроизведения себе подобных (сердце, желудок, печень, селезенка и т. д.);

– исполнительных механизмов (руки, ноги, пальцы, голосовые связки, зубы, органы лица и т. п.);

– управляющего устройства – головного мозга, который принимает решения на основе имеющихся знаний и принимаемых электрических сигналов от первичных преобразователей и формирует управляющие сигналы, передаваемые в исполнительные механизмы.

Как и любая другая органическая материя подвержена износу и старению, что в конце концов приводит к прекращению ее деятельности – смерти.

В результате научных исследований выяснено, что основным функциональным элементом мозга является то, что названо нейроном [6]. Однако вопрос взаимодействия нейронов головного мозга остается проблемным. На основе эмпирического и теоретического опыта можно сделать предположение, что общее устройство головного мозга можно представить в виде

ядра и окружающей его оболочки. Состояние ядра дается человеку от рождения и, скорее всего, не изменяется или подвержено минимальным изменениям. Деятельность ядра головного мозга определяет наши способности и осуществляет верховное управление.

Вторая часть мозга – это самообучаемая нейронная сеть. Она имеет возможность реконфигурироваться, получать и накапливать информацию, производить настройку обрабатывающих функций и коэффициентов передачи для сигналов, поступающих через дендриты в нейрон, формируя тем самым для данного момента времени систему принятия решений. Сущность принятия решений зависит от окружающего пространства и состояния организма, информацию о которых передают в нейронную сеть первичные преобразователи информации. Все действия и ощущения, которые воспроизводит и получает человек, есть сущность и отражение решений, принимаемых нейронной сетью головного мозга. Самообучение нейронной сети производится эмпирически или под целенаправленным воздействием, в том числе и по желанию нейронной сети, т. е. принятых ею решений на проведение определенных действий по получению новой информации.

Если следовать данной модели, то можно аргументировать вывод, что поведение человека обусловливается первоначальной настройкой нейронной сети головного мозга и последующей реконфигурацией самообучаемой, многоуровневой, объемной нейронной сети головного мозга, происходящей по мере получения новой или переработки имеющейся информации.

Работа мозга сопровождается изменением электромагнитного поля, которое можно зафиксировать специальными первичными преобразователями и преобразовать в виде изменяющихся параметров тока или напряжения, что и происходит, когда снимают электроэнцефалограмму. Электрические параметры работа мозга для каждого выполняемого действия свои и они имеют свойство повторяться с допустимым разбросом для каждого действия. Уже достаточно хорошо отработаны вопросы мозгового управления протезами, даже есть разработки, когда человек мысленно управляет движением роботизированного объекта [7].

Новые реакции вырабатываются и запоминаются нейронной сетью на основе либо создания новых синаптических связей между имеющимися нейронами, либо изменения эффективности (коэффициентов передачи и преобразующих функций) уже имеющихся синаптических связей. Запоминание (долговременное) информации – это изменение способности одних нейронов возбуждаться при возбуждении других нейронов.

Если допустить возможность считывания состояния мозга и хранимой в нем информации, а затем ее переноса в другой аналогичный или подобный субъект (объект), то можно представить (определить) данный процесс

как перенос жизни одного человека в другое тело. Совсем не обязательно именно человеческое, а может быть в тело другого какого-то существа, обладающего аналогичными или схожими возможностями, возможно, даже в неорганическую структуру. При этом с большой долей вероятности потребуются перенастройка и дообучение перенесенной настройки и памяти нейронной сети до уровня, необходимого для самообучения и принятия решений новым разумным существом.

В пределе можно допустить как органическое, так и неорганическое или смешанное строение этой новой разумной сущности. В итоге представляется возможным достижение вечной, обогащаемой новыми знаниями жизни. Вполне вероятно, что такое происходит уже сейчас, так как от рождения человеческого тела до его старения и отмирания осуществляется самообучение мозга полученными ранее и сохраненным знаниям, а также создание, накопление и сохранение новых знаний, которые передаются новому поколению вечной жизни. Возможность переноса нейронной сети головного мозга человека в неорганическую структуру предоставляет гораздо большие возможности для переноса жизни на другие планеты и в другие звездные системы, обеспечивая ее неисчезаемость.

Таким образом, разработка систем и средств получения информации о состоянии человека, нейронной сети его головного мозга, об их изменении во времени и под внешним воздействием, является весьма актуальной и позволит решать многие сложные задачи. Построение систем принятия решений, моделирующих отдельные функциональные элементы мозга живых существ, позволит придать системам искусственного интеллекта новые возможности.

Модель базового элемента системы автоматики принятия решений

Состояние наблюдаемого технического, природного и социального объекта описывается параметрами и характеристиками.

Параметр – свойство (показатель) объекта или системы, которое можно измерить. Результатом измерения параметра является число или величина, а сам объект или систему можно рассматривать как множество параметров, которое требуется определить или измерить для моделирования или оценки ее поведения. Иногда параметрами называют также величины, очень медленно изменяющиеся по сравнению с другими величинами (переменными). Примерами параметров вибрации являются масса, температура, амплитуда колебаний, давление, возраст и число жителей.

Характеристика – совокупность отличительных свойств кого-либо или чего-либо. Это графическое или табличное выражение зависимости одного параметра от другого, а также функция, выражающая или описывающая данную зависимость. Например, характеристикой объекта является ампли-

тудный спектр исследуемого сигнала, электрокардиограмма, изменение во времени цены на конкретную продукцию.

Для выполнения оценки состояния наблюдаемого объекта требуется некая система принятия решений или поддержки принятия решений. Предлагается следующая модель базового решающего элемента системы принятия решений по оценке состояния наблюдаемого объекта или выработке рекомендаций по воздействию на данный объект. Такой базовый решающий элемент может служить и моделью нейрона.

Входными данными базового элемента являются:

x_i – значение параметра i, $i = 1,...,N$;

$\omega_j\left(y_{j,1},...,y_{j,k}\right)$ – характеристика j при дискретных значениях аргумента y_j, $j = 1,...,M$;

$\omega_j\left(y_j(t)\right)$ – характеристика j при непрерывном значении аргумента y_j, $j = 1,...,M$.

По отношению к входным исходным параметрам и характеристикам применяются функции первичной обработки:

$$f_l(x_i), \text{ где } l = 1,...,B, \text{ и } \varphi_m\left(\omega_j\right), \text{ где } m = 1,...,C.$$

Причем разные функции $f_l(\ )$ могут применяться по отношению к одному и тому же параметру x_i, а различные функции $\varphi_m(\ )$ – к одному и тому же значению характеристики ω_j. Предполагается применение и комплексных многопараметрических-многохарактеристических функций:

$$\psi_n\left(x_i,...,x_j,...,x_k,\omega_l,...,\omega_m,...,\omega_p\right), \quad \text{где} \quad n = 1,...,D; \quad i,j,k \in 1,...,N; \quad l,m,p \in 1,...,M.$$

По отношению к набору функций: $f_l(x_i)$, $\varphi_m\left(\omega_j\right)$, $\psi_n\left(x_i,...,x_j,...,x_k,\omega_l,...,\omega_m,...,\omega_p\right)$ применяются обобщающие функции:

$$y_k = \Psi_k\lfloor f_l(x_i), l = 1,...,B, \varphi_m\left(\omega_j\right),$$
$$m = 1,...,C, \psi_n\left(x_i,...,x_j,...,x_k,\omega_l,...,\omega_m,...,\omega_p\right), n = 1,...,D\rfloor, \ k = 1,...,L.$$

Уже по отношению к y_k применяются разнообразные решающие функции: $S_\eta\left(y_k\right)$, $\eta = 1,...,P$.

Результат функции $S_\eta(y_k)$ определяет одно из возможных состояний анализируемого объекта, тип самого объекта, принимаемое решение. На рисунке данная модель представлена в графическом виде.

В простейшем случае параметры предлагаемой модели принятия решений будут иметь следующий вид:

$$f_l(x_i) = a_l x_i, \quad \text{где} \quad l = 1,...,B, \quad \varphi_m(\omega_j) = b_m \omega_j, \quad \text{где} \quad m = 1,...,C;$$

a_l, b_m – действительные числа;

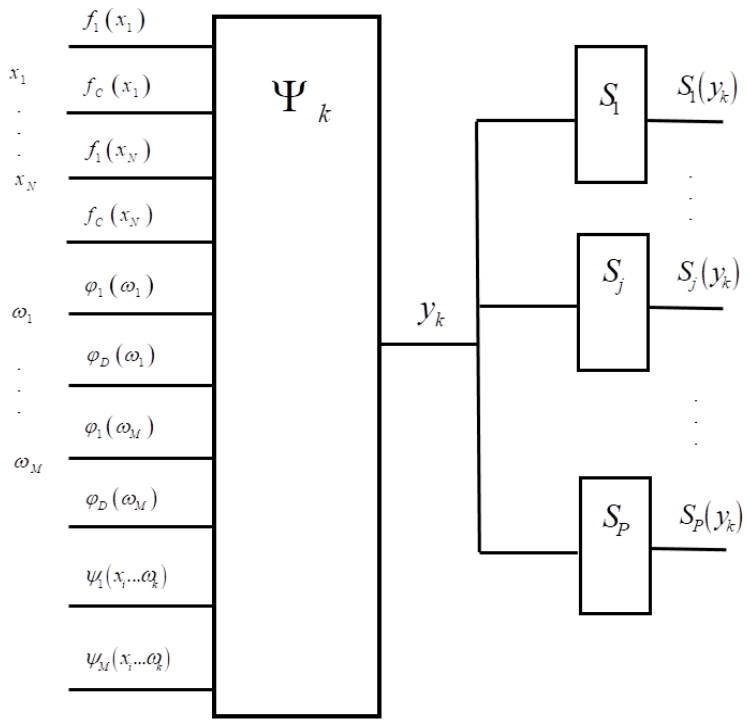
$$\psi_n(x_i,...,x_j,...,x_k,\omega_l,...,\omega_m,...,\omega_p) = c_n\left(\sum_{i=1}^{N} r_i x_i + \sum_{j=1}^{N} s_j \omega_j\right),$$

где $n = 1,...,D$; $i, j, k \in 1,...,N$; $l, m, p \in 1,...,M$; r_i, s_j – действительные числа.

$$y_k = \sum_{l=1}^{B} u_{l,k} f_l(x_i) + \sum_{m=1}^{C} v_{m,k} \varphi_m(\omega_j) + \sum_{n=1}^{D} w_{n,k} \psi_n(x_i,...,x_j,...,x_k,\omega_l,...,\omega_m,...,\omega_p),$$

где $k = 1,...,L$; $u_{l,k}$, $v_{m,k}$, $w_{n,k}$ – действительные числа.

$$S_\eta(y_k) = \rho_\eta y_k, \quad \eta = 1,...,P.$$



Модель базового элемента принятия решений
по оценке состояния наблюдаемого объекта

Пример применения базового элемента системы принятия решений в автоматике защитного отключении турбоагрегата по параметрам вибрации

Наиболее важной задачей современных систем вибрационного контроля и диагностики является предотвращение аварийного повреждения защищаемого объекта при внезапном возникновении неисправностей (механических повреждений) в его узлах или при существенном отклонении каких-либо технологических параметров от номинальных. Особенно это важно для энергетических объектов, задачей которых является непрерывное производство электрической энергии. Стоимость таких объектов очень высока. Поэтому автоматизации принятия решений по оценке их технического состояния, своевременному предупреждению и отключению уделяется большое внимание. В настоящее время разработан и эксплуатируется ряд измерительно-вычислительных комплексов, решающих данную задачу на основе параметров вибрации подшипниковых опор турбоагрегатов, в которых реализован метод базовых элементов принятия решений [8].

Однако факт возникновения ситуации, требующей останова технического объекта, во многих случаях имеет неоднозначное отображение в параметры вибрации. Стандартизованные критерии защиты отражают наиболее общие взаимосвязи, полученные на основе длительного опыта эксплуатации и исследования механизмов с вращательным движением, и далеко не всегда в полной мере могут удовлетворить эксплуатирующий и управляющий персонал.

Системы вибрационного контроля и защиты, построенные на базе компьютерной техники, позволяют реализовать разнообразные и сложные алгоритмы защиты, ориентированные на конкретные типы дефектов и аварийных ситуаций. Это, в свою очередь, позволяет избежать необоснованных срабатываний защитного отключения («ложная тревога») и не допустить «пропуска дефекта» [9].

Реализован и прошел апробацию на ряде турбоагрегатов алгоритм защитного отключения по вибрации, где учитывается несколько факторов:

1. Фактор низкочастотной составляющей вибрации. Под низкочастотной вибрацией понимается среднее квардратическое значение (СКЗ) виброскорости в зоне частоты, равной половине оборотной. Сигнал защитного отключения вырабатывается в том случае, если для любой подшипниковой опоры турбоагрегата возникла следующая ситуация: СКЗ виброскорости низкочастотной вибрации, измеренное для вертикального и поперечно-горизонтального направления любой подшипниковой опоры, на протяжении 4–6 с превышает v мм/с и, притом хотя бы для одного из этих направлений оно на протяжении этого же времени превышает 3 v мм/с. Уровень v определяется типом и рабочими частотами механизма.

2. Фактор оборотной составляющей вибрации. Под оборотной составляющей вибрации понимается СКЗ виброскорости спектральной составляющей с частотой, равной частоте вращения вала (ротора) агрегата:

– величина СКЗ оборотной составляющей. Для каждой подшипниковой опоры и каждого из направлений измерения вибрации устанавливается значение СКЗ виброскорости оборотной составляющей, соответствующее аварийному уровню, который выбирается с учетом конструктивных, функциональных и эксплуатационных особенностей контролируемого механизма. Сигнал защитного отключения вырабатывается в том случае, если в четырех или более точках контроля СКЗ виброскорости оборотной составляющей превысило заданный для соответствующей точки аварийный уровень;

– вектор приращения оборотной составляющей. Для каждой подшипниковой опоры и каждого из направлений измерения вибрации устанавливается значение вектора приращения оборотной составляющей, соответствующее аварийному уровню. Сигнал защитного отключения вырабатывается в том случае, если в четырех или более точках измерений вектор приращения оборотной составляющей превысил заданный для соответствующей точки измерений аварийный уровень.

3. Фактор высокочастотной составляющей вибрации. Под высокочастотной составляющей вибрации понимается СКЗ виброскорости в частотной полосе, нижняя граница которой равна двойной оборотной частоте, а верхняя − верхней границе частотного диапазона, в котором производится вибрационный контроль наблюдаемого механизма. Сигнал защитного отключения вырабатывается в том случае, если для любых двух направлений измерения вибрации для любой подшипниковой опоры высокочастотная вибрация превысила значение аварийного уровня, установленного для данного объекта в течение 3–6 с.

Сигнал на защитное отключение контролируемого механизма вырабатывается в том случае, если он выработан по одному из указанных критериев или по нескольким критериям одновременно.

Для реализации данной системы принятия решения о защитном отключении в качестве входных данных применяется характеристика: $TI_j\left(x_0,...,x_{N-1}\right)$, $j=1,...,M$ − временная реализация вибрационного сигнала; где M − число точек контроля для наблюдаемого турбоагрегата. Для каждой подшипниковой опоры контроль вибрации осуществляется в трех точках-направлениях: вертикальном, поперечно-горизонтальном, осевом; τ − дискретное время, определяющее периодичность получения исходных вибрационных сигналов.

Функция $\varphi_{1,j,\Delta}\lfloor TI_j(x_0,...,x_{N-1}),\tau_\Delta\rfloor$, $j=1,...,M$; $\Delta=1,2,...$, предназначена для вычисления СКЗ оборотной составляющей вибрации $A_{RCV,j}$.

Функция $\varphi_{2,j,\Delta}\lfloor TI_j(x_0,...,x_{N-1}),\tau_\Delta\rfloor$, $j=1,...,M$; $\Delta=1,2,...$, – для вычисления фазы оборотной составляющей вибрации $\Phi_{RCV,j}$.

Функция $\varphi_{3,j,\Delta}\lfloor TI_j(x_0,...,x_{N-1}),\tau_\Delta\rfloor$, $j=1,...,M$; $\Delta=1,2,...$, – для вычисления СКЗ низкочастотной вибрации.

Функция $\varphi_{4,j,\Delta}\lfloor TI_j(x_0,...,x_{N-1}),\tau_\Delta\rfloor$, $j=1,...,M$; $\Delta=1,2,...$, – для вычисления СКЗ высокочастотной вибрации.

Функция $f_{1,j,\Delta}(A_{CCV,j},\Phi_{CCV,j},\tau_{\Delta-1},\tau_\Delta)$, $j=1,...,M$; $\Delta=1,2,...$, – для вычисления вектора приращения оборотной составляющей.

Обобщающие функции первого уровня имеют вид:

$$y_{1,\Delta}=\Psi_{1,1,\Delta}\lfloor\varphi_{3,j,\Delta}(\ );\tau_\Delta;\ j=1,...,M;\Delta=1,2,...\rfloor;$$

$$y_{2,\Delta}=\Psi_{1,2,\Delta}\left\lfloor\varphi_{2,j,\Delta}(\ );\tau_\Delta;\ j=1,...,M;\Delta=1,2,...\right\rfloor;$$

$$y_{3,\Delta}=\Psi_{1,3,\Delta}\left\lfloor f_{1,j,\Delta}(\ );\tau_\Delta;\ j=1,...,M;\Delta=1,2,...\right\rfloor;$$

$$y_{4,\Delta}=\Psi_{1,4,\Delta}\left\lfloor\varphi_{4,j,\Delta}(\ );\tau_\Delta;\ j=1,...,M;\Delta=1,2,...\right\rfloor.$$

Результатом каждой из обобщающих функций первого уровня $\Psi_{1,k,\Delta}(\ )$, $k=1,...,4$, является $y_{k,\Delta}$, которое принимает два значения: ноль или единица. $y_{k,\Delta}$ являются входными параметрами $x_{k,\Delta}$ для обобщающей функции второго уровня:

$$z_\Delta=\Psi_{2,\Delta}(x_{1,\Delta}=y_{1,\Delta},x_{2,\Delta}=y_{2,\Delta},x_{3,\Delta}=y_{3,\Delta},x_{4,\Delta}=y_{4,\Delta}).$$

Результат обобщающей функции второго уровня z_Δ: значение z_Δ=1 соответствует решению о срабатывании защитного отключения, значение z_Δ=0 – нормальному режиму эксплуатации контролируемого объекта.

Заключение

Предложенная модель базового элемента системы принятия решений позволяет конфигурировать различные структуры, предназначенные для оценки состояния сложных технических и природных объектов. Рассмотрена реализация системы принятия решений по оценке состояния многоопорного энергетического агрегата как комбинации базовых решающих

модулей, которая введена в промышленную эксплуатацию [10, 11]. Предлагаемый подход может быть использован при моделировании разнообразных систем, принимающих решения, в том числе и живых организмов, управляющим устройством которых является многоуровневая, многослойная, объемная нейронная сеть, ее типовой элемент имеет вид предложенного базового элемента принятия решений. Выполнен ряд исследований по исследованию электроэнцефалограммы с целью выявления информативно-значимых параметров, позволивших оценивать эффективность применяемых методов и методик лечения нервно-психических заболеваний.

Список использованных источников

1. Неразрушающий контроль : справочник. Т. 7 : под общ. ред. В. В. Клюева. Кн. 2. Вибродиагностика / Ф. Я. Балицкий, А. В. Барков, Н. А. Баркова [и др.]. – М. : Машиностроение, 2005. – 485 с.

2. Bently, D. E. Fundamentals of Rotating Machinery Diagnostics / D. E. Bently, C. N. Hatch, B. Grissom. – Canada : Bently pressurized bearing company, 2002. – 726 p.

3. Бранцевич, П. Ю. Оценка технического состояния механизмов с вращательным движением на основе анализа вибрационных характеристик пусков и выбегов / П. Ю. Бранцевич. – Минск : Четыре четверти, 2021. – 236 с.

4. Барков, А. В. Мониторинг и диагностика роторных машин по вибрации / А. В. Барков, Н. А. Баркова, А. Ю. Азовцев. – СПб. : Изд. центр СПбГМТУ, 2000. – 169 с.

5. Казак, Т. В. Различия в социально-психологических характеристиках управленческих кадров различных уровней руководства в учреждении высшего образования / Т. В. Казак, А. Г. Зенкевич // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня : сб. материалов VIII Междунар. науч.-практ. конф., Минск, 11–12 мая 2022 г. – Минск : Бестпринт, 2022. – С. 285–290.

6. Осовский, С. Нейронные сети для обработки информации / С. Осовский; пер. с польского И. Д. Рудинского. – М. : Финансы и статистика, 2002. – 344 с.

7. Павлова, Л. П. Доминанты деятельного мозга. Системный психофизиологический подход к анализу ЭЭГ / Л. П. Павлова. – СПб. : ИНФОРМ-НАВИГАТОР, 2017. – 442 с.

8. Бранцевич, П. Ю. Принятие решений в задачах, связанных с обработкой вибрационных сигналов / П. Ю. Бранцевич // BIG DATA and Advanced Analytics = BI DATA и анализ высокого уровня : сб. науч. ст. VIII Междунар. науч.-практ. конф., Минск, 11–12 мая 2022 г. – Минск : Бестпринт, 2022. – С. 440–449.

9. Бранцевич, П. Ю. Организация и алгоритмы системы вибрационного контроля и оценки технического состояния турбоагрегатов по вибрационным параметрам / П. Ю. Бранцевич, С. Ф. Костюк, Г. Г. Соболь // Проблемы вибрации, виброналадки, вибромониторинга и диагностики оборудования электрических станций : сб. докладов; под общей ред. А. В. Салимона. – М. : ВТИ, 2003. – С. 25–29.

10. Бранцевич, П. Ю. Большие данные в системах вибрационного контроля, мониторинга, диагностики / П. Ю. Бранцевич, Е. Н. Базылев // Неразрушающий контроль и диагностика. − 2016. − № 3. − С. 28–41.

11. Brancevich, P. Detection of vibration disturbances during the analysis of long realisations of vibration signals / P. Brancevich, Y. Li // 25-th International Congress on

Sound and Vibration, Hiroshima, Japan, 8-12 July 2018. – USA : Curran Associates, Inc., 2018. – P. 2736–2743.

12. Бранцевич, П. Ю. Примеры цифровой обработки электроэнцефалограмм / П. Ю. Бранцевич // Медэлектроника–2022. Средства медицинской электроники и новые медицинские технологии : сб. науч. ст. XIII Междунар. науч.-техн. конф., Минск, 8–9 дек. 2022 г. – Минск : БГУИР, 2022. – С. 314–318.

13. Бранцевич, П. Ю. Цифровая обработка электроэнцефалограмм / П. Ю. Бранцевич // Science and innovation. – Ташкент : LLC «Science and innovation», 2023. – Special iss. 3. – С. 930–934.

14. Бранцевич, П. Ю. Сравнительный анализ электроэнцефалограмм / П. Ю. Бранцевич // BIG DATA и анализ высокого уровня = BIG DATA and Advanced Analytics: сб. науч. ст. IX Междунар. науч.-практ. конф. : в 2 ч., Минск, 17–18 мая 2023 г. / редкол.: В. А. Богуш [и др.]. – Минск : БГУИР, 2023. – Ч. 1. – С. 132–144.

15. Бранцевич, П. Ю. Информативные параметры электроэнцефалограмм / П. Ю. Бранцевич // Прикладные проблемы оптики, информатики, радиофизики и физики конденсированного состояния : материалы Седьмой Междунар. науч.-практ. конф. Минск, 18–19 мая 2031 г. / М-во образования Респ. Беларусь, НИУ «Ин-т приклад. физ. проблем им. А. Н. Севченко» Беларус. гос. ун-та; редкол.: Ю. И. Дудчик (гл. ред.), И. М. Цикман, И. Н. Кольчевская. – Минск : ОДО «Рейплац», 2023. – С. 137–139.

УДК 004.891

Мультимодальный анализ изображений для диагностики аномалий яичников с использованием машинного обучения

В. В. Венгеренко,
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: vengerenko@lsi.bas-net.by

М. М. Лукашевич
Белорусский государственный университет,
Минск

Введение

В настоящее время можно констатировать, что человечество сталкивается с множеством проблем, связанных с состоянием здоровья и физической формой. Значительная часть этих проблем успешно решается благодаря достижениям медицинской науки, однако ряд заболеваний остаются неизлечимыми. Одной из ключевых причин отсутствия эффективных методов лечения является недостаточная развитость систем ранней диагностики [1]. Проблемы со здоровьем яичников, в частности синдром поликистозных яичников (СПЯ), распространены по всему миру, особенно в Азии. Это эндокринное заболевание, которым страдает значительная часть женщин. В настоящее время состояние пациенток характеризуется сочетанием множества симптомов. Эти многочисленные симптомы являются одним из аспектов осложнений заболевания [2, 3]. В литературе описано множество заболеваний яичников: СПЯ, рак, кисты, эндометриоз, перекрут, герминогенные опухоли и первичная яичниковая недостаточность [4]. Всемирная организация здравоохранения опубликовала данные о том, что от 8 до 13 % населения, имеющего яичники, страдает поликистозом яичников, а доля недиагностированных пациентов составляет 70 % [5]. Число пациенток, страдающих СПЯ, стремительно растет. Это связано с тем, что точного медикаментозного или хирургического лечения заболевания не существует. Ранняя диагностика синдрома для точного назначения лекарств является актуальной задачей. Значительная часть исследований в данной области посвящена разработке методов диагностики и лечения с помощью разнообразных материалов и подходов. Среди них можно выделить фитотерапию [6], методы вычислительного интеллекта [7, 8], модели машинного обучения [9–12] и алгоритмы искусственного интеллекта в целом [13–16].

Однако, несмотря на достигнутые успехи, остается актуальной проблема корректного сбора и обработки данных [9, 14]. В клинической практике аномалии яичников представляют собой распространенное явление. Их диагностика и лечение традиционно осуществляются с помощью таких методов, как анализ крови и ультразвуковое исследование. Однако данные подходы обладают рядом ограничений, что стимулировало исследователей к активному внедрению методов машинного обучения (ML) и искусственного интеллекта (AI) для повышения точности и эффективности диагностики. Несмотря на достигнутые успехи, сложность синдромов, связанных с заболеваниями яичников, остается актуальной проблемой, обусловленной высокой вариабельностью симптоматики и этиологии у различных пациентов.

Традиционные методы выявления проблем с яичниками, учитывающие совокупность симптомов и показатели гормональных тестов, позволяют достичь максимальной точности 82,1 % [17]. Методы ML и AI могут быть применены как для результатов анализа крови, гормональных показателей и др., так и непосредственно к изображениям ультразвуковой диагностики. В проекте основной акцент будет сделан на анализ ультразвуковых изображений яичников. Применение методов искусственного интеллекта на отдельных наборах данных позволяет достичь точности порядка 92 % [7, 13, 14, 17]. В 2023 г. A. Alamoudi [et al.] [18] предложили диагностическую модель, основанную на одновременном анализе данных ультразвуковых изображений и клинических показателей (включая анализ крови и другие параметры) с использованием методов глубокого обучения (DL). Авторы достигли точности диагностики, равной 84,81 %. В исследовании были использованы 250 ультразвуковых изображений, подтверждающих наличие синдрома, в то время как остальная часть данных включала изображения, связанные с другими патологиями брюшной полости. В качестве перспективных направлений дальнейших исследований авторы предложили расширение объема выборки и применение различных методов слияния данных для повышения точности модели. В 2024 г. S. Subramani [et al.] [19] реализовали подход на основе глубокого обучения, используя архитектуру YOLOv5. В исследовании проанализированы 400 ультразвуковых изображений, из которых 124 были положительными случаями, а 276 – отрицательными. В качестве рекомендаций для будущих исследований авторы указывают необходимость расширения набора данных, оптимизации гиперпараметров модели, а также необходимость повышения интерпретируемости результатов. S. Vinothini [et al.] [20] представили модель диагностики СПЯ с применением XGBoost, RF, K Nearest Neighbors и SVM. Набор данных включал 541 запись. Авторы отмечают, что повысить корректность модели можно за счет использования разнообразных наборов

данных. J. Manika [et al.] [21] исследовали проблему шумоподавления на ультразвуковых изображениях для увеличения точности диагностической модели. Было использовано двумерное преобразование Фурье в сочетании с моделью сверточной нейронной сети VGG-16. Для кластеризации изображений применялись нейронные сети и XGBoost. Для классификации использовали сверточные нейронные сети, для оценки шумоподавления применяли такие показатели, как RMSE, отношение пикового сигнала к шуму и SSIM. Для выявления СПЯ в яичниках P. Ghadekar [et al.] [22] применяли мультимодальный подход, объединяющий изображения и текстовые данные. Используя VGGNet-16 и XGBoost, гибридная модель показала высокую точность. P. Moral [et al.] [23] представили архитектуру CystNet для диагностики и классификации СПЯ по ультразвуковым изображениям. Этот подход демонстрирует потенциал для предсказания СПЯ или подобных заболеваний с помощью наборов данных ультразвукового исследования. Исследование было ограничено разнообразием данных, что повлияло на обобщаемость модели. В качестве будущего направления авторы рекомендуют расширять наборы данных. A. H. Mridul [et al.] [4] оценили несколько классификаторов, включая рекуррентные нейронные сети (RNN), сверточные нейронный сети (CNN), LSTM, RF, машины опорных векторов (SVM) и ансамблевые методы. В будущем необходимы исследования для повышения эффективности и масштабируемости прогностических моделей. Кроме того, отдельным направлением является разработка мультимодальных моделей машинного обучения, способных работать как с изображениями разной природы (полученными в видимом диапазоне, ультразвуковыми и т. д.), так и с принципиально отличающимися данными (изображениями, табличными данными).

Состояние вопроса в Республике Беларусь и зарубежном государстве-партнере

В Беларуси существует развитая научная база для исследований в области обработки изображений, которая поддерживается как академическими институтами, так и IT-компаниями. Однако анализ показывает, что нет исследований в области анализа такого класса изображений, как изображения яичников. При этом следует отметить, что общие подходы в области анализа изображений не дадут приемлемой точности анализа такого узкого и специфического класса изображений. Требуются разработка специализированных методов и алгоритмов на основе теории искусственного интеллекта в целом и методов машинного обучения в частности, формализованное описание возможных аномалий и классов болезней. Также имеется существенный пробел в части наборов данных по данной тематике, формулировки и реализации требований для сбора такого класса изображений,

их разметки. Как результат отсутствуют программные средства анализа таких изображений, которые эксплуатировались бы в медицинских учреждениях.

Пакистанская сторона отмечает, что у них существует высокий риск такого рода заболеваний и ограниченный доступ к ранней диагностике, а также проблемы со сбором данных для разработки методов машинного обучения с целью анализа изображений. Отсутствует автоматический анализ таких изображений. Кроме того, на азиатском континенте уровень распространенности аномалий яичников резко возрастает по сравнению с другими развитыми континентами и количество пациентов стремительно растет. Существует острая необходимость разработки методов и средств анализа аномалий на изображениях яичников. Отмечается необходимость разработки мультимодальных моделей, способных работать с разными типами данных.

В настоящее время автоматизированные методы анализа аномалий на изображениях яичников не используются в странах – участницах проекта.

На основании проведенного анализа можно сделать выводы о том, что ряд исследователей сосредоточились на разработке методов обработки изображений для диагностики заболевания [25, 26]. Однако, несмотря на достижение высокой точности в отдельных исследованиях, большинство работ указывают на необходимость дальнейшего совершенствования подходов с использованием разнообразных наборов данных. Это обусловлено рядом критических проблем:

1. *Сложность моделирования поликистозного синдрома.* Этиология и проявления синдрома носят многофакторный характер, что затрудняет построение универсальных моделей. Кроме того, доступные наборы данных часто несбалансированы и неполны, поскольку одни исследования основываются на визуальных данных (например, ультразвуковых изображениях), а другие – на клинических наблюдениях.

2. *Проблема обобщаемости диагностических моделей.* Существующие модели демонстрируют ограниченную применимость к различным популяциям, что может быть связано с неоптимальным выбором признаков для прогнозирования, а также с различиями в наборах данных, используемых в разных регионах мира.

3. *Недостаток доверия со стороны медицинских специалистов.* Связан с отсутствием методов объяснимого искусственного интеллекта (Explainable AI, XAI), которые могли бы повысить прозрачность и интерпретируемость результатов, что особенно важно в медицинской практике. XAI – модель, которая могла бы в перспективе объяснять механизмы, лежащие за алгоритмами машинного обучения.

4. *Отсутствие стандартизированных протоколов сбора данных.* Ограничивает возможность сравнения и интеграции данных из различных источников, что критически важно для создания надежных и воспроизводимых моделей.

Для решения указанных проблем, в частности повышения обобщаемости моделей, требуется разработка комплексных подходов, включающих стандартизацию данных, внедрение методов объяснимого ИИ и создание многофакторных моделей, учитывающих разнообразие клинических и демографических характеристик пациентов. Совокупность указанных проблем четко показывает целесообразность совместных исследований, где белорусская сторона сможет успешно применить уже накопленный опыт в области обработки в конкретной прикладной и не до конца проработанной области, а пакистанская сторона сделает акцент на сборе данных среди популяции, где особенно остро стоят такого рода заболевания.

Таким образом, цель проекта – разработать и исследовать новые методы мультимодального анализа изображений с использованием машинного обучения, ориентированные на автоматизацию и увеличение точности диагностики аномалий яичников.

Для достижения поставленной цели в работе планируется решить следующие основные задачи:

– подготовить вариабельный набор изображений яичников для построения моделей машинного обучения из доступных источников и при взаимодействии с медицинскими учреждениями;

– определить области интереса и классы объектов на изображениях, являющиеся аномальными, выделить их отличительные признаки;

– разработать мультимодальный подход и модель машинного обучения для выделения и классификации аномалий яичников на изображениях;

– разработать модель машинного обучения на основе интерпретируемых подходов искусственного интеллекта (классических методов машинного обучения и визуализации);

– разработать экспериментальные программные средства для анализа аномалий на изображениях яичников.

Исследования будут базироваться на данных, собранных в 20 медицинских центрах Китая, из базы данных, содержащей около 10 000 изображений (URL: https://www.imrpress.com/journal/CEOG/49/4/10.31083/j.ceog4904096/htm), а также на данных, собранных медиками с пакистанской стороны.

Заключение

Результатом выполнения проекта станет разработка принципиально новых методов и алгоритмов, позволяющих выделять аномалии на изображениях яичников. Для этого будет создана мультимодальная модель,

работающая с разнородными типами данных, а также модель, позволяющая выполнить и интерпретировать результаты анализа изображений. Это повысит доверие к результатам автоматического анализа. Будут разработаны экспериментальные программные средства анализа изображений яичников, что позволит проводить эффективную раннюю диагностику заболеваний.

Полученные в процессе выполнения проекта результаты будут использованы для разработки методики и программных средств анализа изображений яичников. Социальная эффективность проекта заключается в создании новых, объективных методов анализа изображений яичников для ранней диагностики аномалий и заболеваний. Разрабатываемые методы будут описаны в виде алгоритмов и реализованы как программно- и аппаратно-независимые модули, которые можно будет использовать в экспертных организациях.

Сотрудничество между Беларусью и Пакистаном повысит технологические компетенции в области искусственного интеллекта, машинного и глубокого обучения и улучшит социально-экономические отношения между двумя странами.

Список использованных источников

1. Health is Wealth: Meaning, Examples, Origin, Expansion, Significance and Short Stories – Write A Topic. – URL: https://writeatopic.com/proverbs/health-is-wealth-proverb (date of access: 30.10.2024).

2. A Review on the Detection Techniques of Polycystic Ovary Syndrome Using Machine Learning / S. Ahmed, Md. Sazzadur Rahman, Ismate Jahan, M. Shamim Kaiser // IEEE Access. – 2023. – Vol. 11. – P. 86522–86543. – DOI: 10.1109/ACCESS.2023.3304536.

3. Comparison of physical activity levels and dietary habits between women with polycystic ovarian syndrome and healthy controls of reproductive age: a case-control study / M. S. Butt, J. Saleem, R. Zakar [et al.] // BMC Women's Health. – 2024. – Vol. 24, iss. 1. – 10 p. – DOI: 10.1186/s12905-023-02866-3.

4. Mridul, A. H. Polycystic Ovary Syndrome (PCOS) Disease Prediction Using Traditional Machine Learning and Deep Learning Algorithms / A. H. Mridul [et al.] // International Journal of Computer and Information Systems and Industrial Management Applications. – 2024. – Vol. 16, iss. 3. – 25 p.

5. Polycystic ovary syndrome. – URL: https://www.who.int/news-room/fact-sheets/detail/polycystic-ovary-syndrome (date of access: 29.10.2024).

6. Attri, M. Nature’s Remedy: Exploring Herbal Treatments for Polycystic Ovarian Syndrome / M. Attri, A. Raghav, H. Chaudhary // Current Nutraceuticals. – 2024. – Vol. 5. – 17 p.

7. F-Net: Follicles Net an efficient tool for the diagnosis of polycystic ovarian syndrome using deep learning techniques / Sowmiya S, Snekhalatha Umapathy, Omar Alhajlah [et al.] // PLOS ONE. – 2024. – Vol. 19, iss. 8. – 22 p. – DOI: 10.1371/journal.pone.0307571.

8. Computational intelligence for early detection of infertility in women / R. Subha, Radhakrishnan Rekha, M. Selvadass [et al.] // Engineering Applications of Artificial Intelligence. – 2024. – Vol. 127, iss. PB. – 14 p. – DOI: 10.1016/j.engappai.2023.107400.

9. Modi, N. Detection and Classification of Polycystic Ovary Syndrome using Machine Learning-Based Approaches / N. Modi, Y. Kumar // 2024 IEEE Intern. Conf. on Interdisciplinary Approaches in Technology and Management for Social Innovation (IATMSI). – Gwalior, India, 2024. – P. 1–6. – DOI: 10.1109/IATMSI60426.2024.10503222.

10. Predicting polycystic ovary syndrome with machine learning algorithms from electronic health records / Z. Zad, V. S Jiang, Amber T Wolf [et al.] // Frontiers in Endocrinology. – 2024. – Vol. 15. – 14 p. – DOI: 10.3389/fendo.2024.1298628.

11. PCOS Detection and Monitoring using Machine Learning / B. Nimmala, U. D. Nimmala, A. Elangi, S. Bagade // 2024 5th Intern. Conf. on Image Processing and Capsule Networks (ICIPCN). – Dhulikhel, Nepal, 2024. – P. 238–242. – DOI:10.1109/ICIPCN63822.2024. 00046.

12. Satpute, S. A Systematic Review on: Polycystic Ovary Syndrome Using Machine Learning / S. Satpute // International Journal of Research in Applied Science & Engineering Technology. – 2024. – Vol. 12, iss. 1. – P. 162–165. – DOI: 10.22214/ijraset.2024.57876.

13. Unveiling the Role of Artificial Intelligence (AI) in Polycystic Ovary Syndrome (PCOS) Diagnosis: A Comprehensive Review / P. Verma, P. Maan, R. Gautam, T. Arora // Reproductive Sciences. – 2024. – Vol. 31, iss. 10. – P. 2901–2915. – DOI: 10.1007/s43032-024-01615-7.

14. Survey of AI-driven techniques for ovarian cancer detection: state-of-the-art methods and open challenges / S. Singh, M. Kumari Maurya, N. Pratap Singh, R. Kumar // Network Modeling Analysis in Health Informatics and Bioinformatics. – 2024. – Vol. 13, iss. 1. – 56 p. – DOI: 10.1007/s13721-024-00491-0.

15. Tsaliki, K. C. AI-Driven Hormonal Profiling: A Game-Changer in Polycystic Ovary Syndrome Prevention / K. C. Tsaliki // International Journal of Research in Applied Science & Engineering Technology. – 2024. – Vol. 12, iss. 5. – P. 364–371. – DOI: 10.22214/ijraset.2024.61001.

16. Pallavi, C. S. Literature Review on Early PCOS Detection on Girl Child Using Artificial Intelligence or Machine Learning / C. S. Pallavi, S. Soumya, V. C. Narasanakuppe // Sparklinglight Transactions on Artificial Intelligence and Quantum Computing. – 2024. – Vol. 4, iss. 1. – P. 17–31. – DOI: 10.55011/STAIQC.2024.4102.

17. Diagnostic Accuracy of Anti-Mullerian Hormone for Polycystic Ovary Syndrome / K. Bahadur, A. Ijaz, S. Baqai, N. Asif // Pakistan Armed Forces Medical Journal. – 2023. – Vol. 73, iss. 2. – P. 329–332. – DOI: 10.51253/pafmj.v73i2.4249.

18. A Deep Learning Fusion Approach to Diagnosis the Polycystic Ovary Syndrome (PCOS) / A. Alamoudi, I. Ullah Khan, N. Aslam, N. Alqahtani // Applied Computational Intelligence and Soft Computing. – 2023. – Vol. 2023. – 15 p. – DOI: 10.1155/2023/9686697.

19. Subramani, S. Utilizing deep learning techniques for detection of Polycystic Ovarian Syndrome using ovarian ultrasound image / S. Subramani, A. Rarichan, S. H. Chaithra // 2023 Intern. Conf. on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES). – Chennai, India, 2023. – P. 1–7. – DOI: 10.1109/ICSES60034.2023.10465547.

20. Vinothini, S. Polycystic Ovary Syndrome (PCOS) Disease Prediction Using Machine Learning / S. Vinothini, S. Vaishnavi, N. Mythili // 2024 IEEE Intern. Conf. on Contemporary Computing and Communications (InC4). – Bangalore, India, 2024. – P. 1–9. – DOI: 10.1109/InC460750.2024.10649123.

21. Jha, M. Noise cancellation of polycystic ovarian syndrome ultrasound images using robust two-dimensional fractional fourier transform filter and VGG-16 model / M. Jha,

R. Gupta, R. Saxena // International Journal of Information Technology. – 2024. – Vol. 16, iss. 4. – P. 2497–2504. – DOI: 10.1007/s41870-024-01773-6.

22. Multimodal PCOS Detection: Combining XG Boost for Images with Zero Shot Learning for Textual Data / P. Ghadekar, S. Khare, D. Sakharwade [et al.] // 2024 Asia Pacific Conf. on Innovation in Technology (APCIT). – Mysore, India, 2024. – P. 1–8. – DOI: 10.1109/APCIT62007.2024.10673443.

23. Moral, P. PODBoost: an explainable AI model for polycystic ovarian syndrome detection using grey wolf-based feature selection approach / P. Moral, D. Mustafi, S. K. Sahana // Neural Computing and Applications. – 2024. – Vol. 36, iss. 30. – P. 18627–18644. – DOI: 10.1007/s00521-024-10171-9.

УДК 339.138

Интеллектуальные мультиагентные системы

Т. В. Владыкина
Академия управления при Президенте
Республики Беларусь, Минск
E-mail: tatsianavl@mail.ru

Ю. А. Постарнакова
Информационно-издательский центр
по налогам и сборам, Минск, Беларусь

Введение

Прогресс в современном мире неразрывно связан с развитием интеллектуальных информационных систем (ИИС) и мультиагентных технологий, обеспечивающих обработку больших данных и принятие эффективных решений в условиях повышенной динамики и неопределенности. Исследование и использование таких систем приобретают особую значимость для широкого спектра отраслей, включая экономику, производство, образование и здравоохранение, становясь ключевым фактором успеха в современной высокотехнологичной среде.

Понятие и эволюция ИИС

ИИС играют важную роль, сосредоточив в себе наиболее наукоемкие технологии с высоким уровнем автоматизации не только процессов подготовки информации для принятия решений, но и самих процессов выработки вариантов решений, опирающихся на полученные информационной системой данные. ИИС находят широкое применение в таких областях, как анализ деятельности предприятия, стратегическое планирование, инвестиционный анализ, оценка рисков и формирование портфеля ценных бумаг, финансовый анализ, маркетинг и т. д.

Возникновение ИИС является закономерным этапом эволюции классических информационных систем, широко применяемых на предприятиях.

Информационные технологии, применяемые в рамках систем реинжиниринга бизнес-процессов (Business Process Reengineering, BPR), продемонстрировали свою способность эффективно структурировать потоки информации и оптимизировать организационную структуру предприятий. Они способствовали освоению методик построения корпоративных информационных моделей, а также обеспечили комплексную информацион-

ную поддержку всех ключевых аспектов производственной деятельности предприятий, охватывая процессы проектирования, изготовления, сбыта продукции, финансового анализа, кадрового менеджмента, маркетинга, сопровождения выпускаемых товаров и долгосрочного планирования.

Внедрение масштабируемых информационных систем класса ERP (Enterprise Resource Planning) позволило повысить производительность предприятий, создав основу для полной цифровизации межкорпоративных взаимодействий (B2B) и взаимодействия с конечными пользователями (B2C).

Широкая практика успешного внедрения интегрированных информационных систем с фиксированными алгоритмами бухгалтерского учета, финансового анализа, документооборота, производственного планирования и управления запасами привела к росту интереса среди подразделений, занимающихся перспективным развитием, к потенциалу новых возможностей, предоставляемых такими системами. Это обусловило потребность в существенно большем объеме информации как внутренней, касающейся самого предприятия, так и внешней, связанной с окружающей средой, экономическими, политическими факторами, конкурентами, поставщиками и прочими заинтересованными сторонами. Дополнительно возникла необходимость обработки сложных вычислительных операций, включения трудноформализуемых элементов, повышения качества пользовательского интерфейса.

Такие требования были удовлетворены посредством появления специализированных систем поддержки принятия решений (Decision Support Systems, DSS), характеризующихся высоким уровнем интеллектуальной составляющей по сравнению с традиционными системами управления предприятием. Отличительными чертами DSS являются: развитые механизмы сбора и фильтрации данных, регулярный импорт сведений из сторонних источников, предварительный агрегирующий расчет заранее востребованных показателей, специализированная организация хранения данных, позволяющая проводить многомерный анализ различной глубины детализации. Эта технология известна как концепция хранилищ и витрин данных совместно с технологией оперативной аналитической обработки данных (Online Analytical Processing, OLAP).

Для реализации полноценного функционала ИИС необходимы дополнительные технологические компоненты, включающие инструменты управления базами знаний на основе объектно-ориентированного подхода, автоматизированные процедуры наполнения баз знаний методами интеллектуального анализа данных, полнотекстовые поисковые системы и семантические анализаторы естественного языка для организации удобного интерфейсного взаимодействия пользователей с системой.

Основные характеристики ИИС

Эффективность ИИС обусловлена рядом важных признаков:

– развитые коммуникативные способности (возможность обработки произвольных запросов в диалоге на языке, максимально приближенном к естественному);

– направленность на решение слабоструктурированных, плохо формализуемых задач;

– способность работать с неопределенными и динамичными данными;

– способность к развитию системы и извлечению знаний из накопленного опыта конкретных ситуаций;

– возможность получения и использования информации, которая явно не хранится, а выводится из имеющихся в базе данных;

– система имеет не только модель предметной области, но и модель самой себя, что позволяет ей определять границы своей компетентности;

– способность к аддуктивным выводам, т. е. выводам по аналогии;

– способность объяснять свои действия, неудачи пользователя, предупреждать пользователя о некоторых ситуациях, приводящих к нарушению целостности данных [1, с. 28].

Таким образом, ИИС обладают всеми признаками автоматизированных помощников высокого уровня, способных поддерживать аналитическое мышление пользователей и предлагать обоснованные рекомендации для принятия решений.

Понятие и характеристики мультиагентных систем

Мультиагентные системы (МАС) заимствуют идеи из биологии и социальной активности человечества. Основной принцип заключается в формировании ансамбля агентов, совокупная сила которого превышает сумму усилий каждого отдельного элемента.

МАС – это технологии, объединяющие в себе технологии искусственного интеллекта и методы самоорганизации, которые позволяют более эффективно решать многие задачи.

Важнейшей характеристикой МАС является автономия, варьирующая от полной зависимости до абсолютной свободы действий агентов.

Типичным примером является ситуация, когда задача делится на мелкие компоненты, распределяемые среди множества исполнителей. Эффективность МАС зависит от способности агентов взаимодействовать и согласовывать свои действия.

Масштабировать агентный подход к более крупным и сложным системам можно двумя способами: увеличивать мощность агента, чтобы он мог справиться с более ресурсоемкими задачами, либо увеличивать число агентов [2, с. 13].

Многоагентный подход можно реализовать только при выполнении следующих условий: сепарабельность задачи и способность агентов координировать свои действия. Сепарабельность задачи – это принципиальная возможность разделить ее на отдельные задачи, выполнять которые можно независимо друг от друга. Второе условие реализуемости мультиагентного подхода – возможность и желание агентов координировать свои действия [2, с. 16].

Помимо реализуемости следует оценивать целесообразность использования множества агентов. Естественная причина объединения агентов – невозможность выполнить задачу силами одного агента, физически или в заданное время.

Классифицирующим признаком МАС выступает степень автономности агентов. При распределенном подходе автономность отсутствует: все агенты управляются централизованно и каждый решает свою задачу, назначенную из центра. Координация действий агентов также осуществляется централизованно. С одной стороны, такая организация максимально усложняет управление и увеличивает связанные с ним издержки, с другой – позволяет обеспечить контролируемое качество.

При децентрализованном подходе агенты могут связываться друг с другом напрямую и решать локальные проблемы, не обращаясь к вышестоящим координаторам. В качестве примера можно привести вошедшие в нашу обыденность роботы-курьеры, которые застревают в снегу так, что случайным прохожим приходится им помогать. Недостатком децентрализованных агентов искусственного интеллекта может быть отклонение от стандартов, принятых в организации. Например, два агента-программиста могут оперативно решить вопрос взаимодействия разрабатываемых ими модулей, что породит недокументированные возможности и приведет к уязвимости системы в целом.

Разделение задачи на подзадачи для поручения их решения множеству исполнителей может осуществляться двумя способами: на основе разделения функций и на основе разделения пространства. В первом случае исполнители должны быть специализированными: каждый исполнитель способен выполнять только одну функцию, а поскольку решение задачи обычно включает в себя множество функций, то исполнители образуют конвейер. Если исполнители задачи делят не функции, а фронт работ, то они могут обладать самостоятельностью в выборе целей и должны координировать свои действия.

В зависимости от совместимости целей, потребности в чужом опыте, наличии распределенных ресурсов существуют следующие типы взаимодействия агентов:

– координируемое сотрудничество;

– простое сотрудничество;
– непродуктивное сотрудничество;
– независимость;
– коллективное соперничество за ресурсы;
– командное сотрудничество;
– индивидуальное соперничество за ресурсы;
– индивидуальное соперничество.

Простое сотрудничество возникает, когда агенты действуют независимо друг от друга, но их цели совместимы. Агенты не конкурируют за ресурсы и могут пользоваться чужим опытом.

Непродуктивное сотрудничество имеет место, когда цель общая, но нет потребности в опыте других участников и, как следствие, нет координации.

Независимость агентов означает только совместимость целей. Агенты не конкурируют за ресурсы и не нуждаются в обмене опытом.

Коллективное соперничество за ресурсы возникает, если цели индивидуальные, опыт их достижения отсутствует, а ресурсов одного агента недостаточно для достижения цели. В таком случае агенты вынуждены создавать группы для совместного решения задачи, отвечающей цели участников.

Командное сотрудничество – это тип взаимоотношений, базирующийся главным образом на необходимости обмена опытом, т. е. на координации усилий. Члены одной команды не конкурируют за ресурсы, а каждый может преследовать собственные цели.

Индивидуальное соперничество за ресурсы появляется, если есть распределенные ресурсы, но участников процесса не объединяют ни совместные цели, ни необходимость обмена опытом.

Индивидуальное соперничество возникает, если участники группы не имеют не только совместных целей и обмена опытом, но и необходимости бороться за распределенные ресурсы нет. Появляется чистое соперничество, мотивация которого лежит вне сферы борьбы за выживание.

Таким образом, объединение агентов в ансамбль может быть принудительным, в виде организации команды, или добровольным, в инициативном порядке. В любом случае требуется взаимодействие агентов, роль которого увеличивается по мере возрастания автономности и уменьшается, если агентов не объединяет общая или совместимая цель [3, с. 21–25].

Проектирование ИИС и МАС

Проектирование ИИС незначительно отличается от технологии проектирования информационных систем. Основная специфика связана с разработкой базы знаний.

ИИС особенно эффективны в применении к слабоструктурированным задачам, в которых отсутствует строгая формализация и для решения которых применяются эвристические процедуры, позволяющие в большинстве случаев получить решение.

Включение в состав ИИС классических экономико-математических моделей, методов линейного, квадратичного и динамического программирования позволяет сочетать анализ объекта на основе экономических показателей с учетом факторов и рисков политических и внеэкономических показателей, оценивать последствия полученных ИИС решений.

Наличие в составе ИИС объектно-ориентированной базы данных позволяет однородными средствами обеспечить хранение и актуализацию как фактов, так и знаний [1, с. 29].

Наиболее распространенные техники реализации МАС включают микросервисную архитектуру и распределенные вычисления.

В микросервисной архитектуре задача разбивается на отдельные функции, которые оформляются в виде отдельных программ – микросервисов. Микросервисы могут быть пассивными агентами, если они запускаются по команде сверху, или активными, если они запускаются при появлении входных данных. Распределенные вычисления – это способ распараллеливания вычислений для их ускорения и эффективного использования вычислительных ресурсов, в частности процессорных ядер. Агенты (отдельные процессы) работают параллельно и могут конкурировать за ресурсы, а также требуют координации использования общей памяти.

Разработка МАС является нетривиальной задачей, которую сложно выполнить без компьютерного моделирования. Существуют разнообразные инструментальные средства разработки и моделирования МАС. Среди популярных инструментов выделяется платформа JADE (Java Agent Development Framework), которая поддерживает создание распределенных и автономных агентов, обеспечивая надежность и безопасность коммуникации между ними. Кроме того, универсальным и удобным инструментом моделирования МАС является система моделирования Anylogic [2, с. 28].

Преимущества и ограничения МАС

Одним из важнейших достоинств МАС является их масштабируемость. Благодаря модульной структуре и поддержке добавления новых агентов МАС демонстрируют высокую гибкость в условиях растущих объемов задач и потребностей предприятий.

Второе значимое преимущество – устойчивость, достигаемая за счет распределения нагрузки и отсутствия единой точки отказа. Выход из строя одного агента не нарушает функционирование всей системы, что крайне

важно для областей с повышенными требованиями к надежности, таких как финансовая сфера и медицина. Также важна специализация агентов, позволяющая каждому элементу заниматься решением определенного круга задач, что улучшает качество итогового результата. Кроме того, одной из значимых особенностей МАС является их способность гармонично вписываться в современные бизнес-процессы. В отличие от других инновационных технологий внедрение МАС не требует кардинальных преобразований существующих инфраструктур. Такие системы могут постепенно интегрироваться в действующие рабочие процессы, улучшая их эффективность без существенных затрат и временных потерь. Благодаря своим уникальным характеристикам МАС создают предпосылки для формирования принципиально нового типа экономики – цифровой экономики искусственного интеллекта, в которой искусственный интеллект играет определяющую роль [4].

Вместе с тем в области МАС есть несколько ключевых проблемных областей и вызовов, связанных с масштабируемостью, интеграцией и адаптивностью. Вызовы представляют собой важные аспекты, которые требуют внимательного рассмотрения и решения для достижения успешной реализации и эксплуатации таких систем.

Необходимость решения проблем масштабируемости, интеграции и адаптивности обусловлена тем, что МАС часто применяются в критически важных и сложных областях, где любые сбои могут иметь значительные последствия.

Для *масштабируемости* необходимы методы и инструменты, которые позволяют системам эффективно справляться с увеличением нагрузки и числа агентов. Может включать использование распределенных вычислений и облачных технологий для обработки больших объемов данных, а также разработку алгоритмов, которые минимизируют потребление ресурсов и обеспечивают высокую производительность при увеличении масштабов системы. Оптимизация архитектуры и внедрение динамического управления ресурсами играют ключевую роль в поддержке масштабируемости.

В области *интеграции* критически важно разработать стандарты и протоколы, которые обеспечивают совместимость и эффективное взаимодействие между различными компонентами системы. Включает в себя создание унифицированных интерфейсов для обмена данными взаимодействия между агентами и интеграцию с существующими системами и платформами.

Поддержка различных форматов данных и протоколов связи, а также обеспечение безопасности и надежности при интеграции с внешними системами являются важными аспектами для успешной реализации мультиагентных систем.

Адаптивность в свою очередь требует разработки методов, которые позволяют агентам динамически изменять свои стратегии и поведение в ответ на изменения в окружающей среде и условиях работы. Может включать внедрение механизмов самообучения и самооптимизации, которые позволяют агентам автоматически корректировать свои действия на основе новых данных и опыта. Управление неопределенностью и адаптация к непредсказуемым условиям являются важными аспектами для обеспечения гибкости и устойчивости системы [2, с. 121–122].

Перспективы развития МАС

Перспективные направления исследований и развивающиеся области МАС представляют собой область с большим потенциалом для инноваций и улучшений. Одним из ключевых направлений является развитие методов и алгоритмов для повышения автономности агентов, что позволяет создавать более интеллектуальные и адаптивные системы.

Современные исследования сосредоточены на создании агентов, способных самостоятельно обучаться и адаптироваться к новым условиям без необходимости постоянного вмешательства со стороны человека.

Другим важным направлением является развитие технологий для обеспечения безопасности и конфиденциальности в распределенных системах. Поскольку МАС часто работают в открытых потенциально небезопасных средах, создание методов защиты данных и предотвращения атак становится критически важным. Исследования в этой области направлены на разработку надежных механизмов шифрования, аутентификации и защиты от несанкционированного доступа.

Еще одним перспективным направлением является интеграция МАС с новыми вычислительными технологиями, такими как edge computing и облачные вычисления. Исследования в этой области сосредоточены на разработке методов распределенного управления и координации агентов, которые могут эффективно использовать возможности edge computing для улучшения производительности и надежности систем.

Не менее важным является развитие МАС для анализа больших данных и интернета вещей. В условиях стремительного роста объемов данных и подключенных устройств МАС могут играть ключевую роль в обработке и анализе информации, а также в управлении и координации действий множества устройств. Исследования в этой области направлены на создание методов и инструментов, которые позволяют эффективно обрабатывать и использовать данные, поступающие от различных источников, и обеспечивать интеллектуальное управление на основе анализа таких данных.

Развитие технологий и методов в этих направлениях открывает новые возможности для создания более умных, безопасных и эффективных систем, которые смогут решать сложные задачи и адаптироваться к постоянно изменяющимся условиям [2, с. 120–121].

Заключение

МАС представляют собой важный этап эволюции ИИС. Несмотря на очевидные преимущества масштабируемости, устойчивости и интеллектуальной обработки данных, все еще не до конца остаются решенными вопросы координации агентов, конфликтующих решений и безопасности. Тем не менее несмотря на существующие ограничения, благодаря широкому спектру возможных применений и активным исследованиям, направленным на устранение текущих недостатков, представляется вероятным, что МАС займут лидирующую позицию в последующих поколениях информационных технологий.

Будущее МАС связано с дальнейшими инновациями в области вычислительных технологий, алгоритмов и взаимодействия. Продолжение исследований и разработок в этой области откроет новые возможности для создания более эффективных, интеллектуальных и автономных систем, которые будут способны справляться с комплексными задачами и требованиями современного мира [2, с. 134].

Список использованных источников

1. Романов, В. П. Интеллектуальные информационные системы в экономике: учеб. пособие / В. П. Романов; под ред. д.э.н., проф. Н. П. Тихомирова. – М. : Издательство «Экзамен», 2003 – 496 с.

2. Проектирование мультиагентных систем: стратегическое направление в искусственном интеллекте : учебник для вузов / А. В. Золкин, Р. А. Вербицкий, С. С. Мясников [и др.]. – СПб. : Лань, 2025. – 160 с.

3. Бессмертный, И. А. Искусственный интеллект. Введение в многоагентные системы : учебник для вузов / И. А. Бессмертный. – М. : Издательство «Юрайт», 2025. – 148 с.

4. Как мультиагентные архитектуры изменят ИИ-ландшафт [сайт] // ООО «Издательство СК Пресс», 2024. – URL: https://www.itweek.ru/ai/article/detail.php?ID=230873 (дата обращения: 30.09.2025).

УДК 004.056:004.8:005.8

Методы защиты информационных систем, использующих Large Language Models

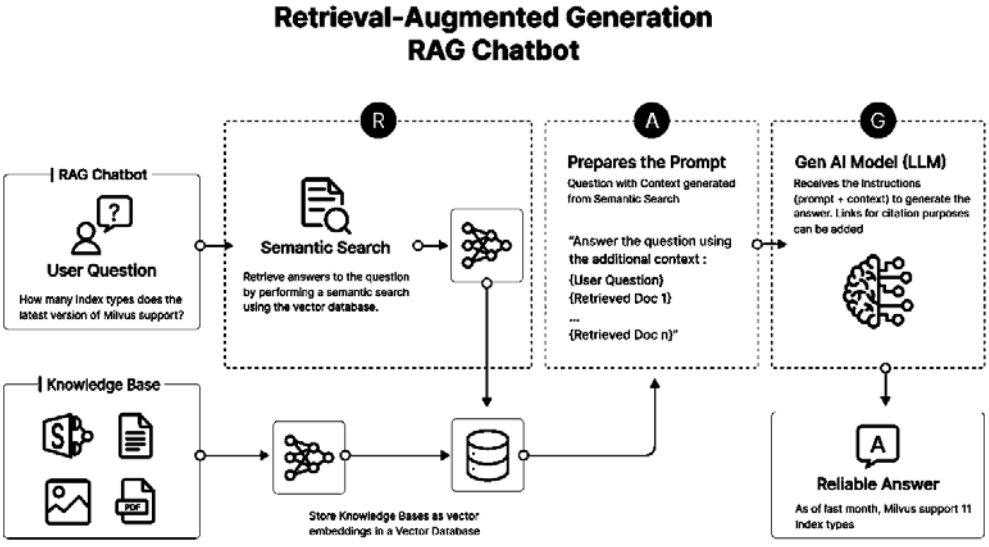
В. Д. Владымцев
ООО «СТЭКЛЭВЭЛ ГРУПП», Минск, Беларусь

Введение

Нарастающая интеграция больших языковых моделей (Large Language Models, LLM) в бизнес-процессы и госуслуги резко расширила поверхность атаки и вывела на первый план специфические уязвимости этих систем [1]. Современные подходы к защите строятся как оборона «в глубину» и покрывают четыре взаимосвязанных слоя: данные, модель, инфраструктуру и этические/комплаенс-контроли, включающие в себя контент-безопасность, локализацию политики и правовые требования [2]. Запрос, по которому LLM генерирует ответ, называется промптом (от англ. prompt – подсказка).

Обобщенная архитектура LLM-системы

Рассмотрим базовую архитектуру информационной системы, использующую LLM в основе ядра приложения. Как правило, система включает фронтенд (например, чат-интерфейс или API), собственно модель LLM (развернутую локально или вызываемую через облачный API), а также внешние модули для расширения возможностей модели.



Базовая архитектура LLM-приложения с Retrieval-Augmented Generation

Одной из распространенных схем является Retrieval-Augmented Generation (RAG) – сочетание LLM с векторной базой знаний для хранения доменных данных. В такой системе пользовательский запрос сначала используется для поиска релевантных фрагментов информации в векторном хранилище, после чего модель генерирует ответ с учетом найденного контекста [3]. На рисунке показана упрощенная RAG-архитектура: LLM дополняется векторной базой, которая обеспечивает поиск знаний и снабжает модель актуальными данными, минимизируя галлюцинации модели за счет более релевантного контекста [3].

Текущие средства информационной безопасности

При проектировании LLM разработчики уже применяют некоторые средства защиты по умолчанию. Во-первых, провайдеры LLM внедряют фильтрацию контента: например, OpenAI использует модерационные правила и классификаторы, которые пытаются блокировать ответы с токсичным или запрещенным содержанием. Тем не менее эти фильтры не гарантируют абсолютной защиты – исследования демонстрируют, что злоумышленники регулярно находят обходные пути [2]. Во-вторых, практикуется разделение ролей в подсказках: системный промпт задает модели ограничения, в том числе, что она «не должна делать», пользовательский ввод идет отдельно [4]. Также рекомендуется ограничивать длину и формат вводимых данных, использовать проверку ввода как и в классических веб-приложениях, чтобы снизить риск промпт-инъекций и других атак [5]. Третьим слоем безопасности является инфраструктурный: развертывание LLM-модели в изолированном окружении (контейнеры, песочницы) и ограничение ее доступа к ресурсам системы. Например, если LLM-сервис запускается с минимальными привилегиями и не имеет прямого доступа к внутренним базам, последствия компрометации будут локализованы. Для RAG-систем внедряются механизмы аутентификации и авторизации на этапах загрузки и поиска данных: каждое обращение к векторной базе данных может требовать API-ключ или OAuth, а выдача данных – проверять права конкретного пользователя [3]. С 2024 г. организации начали внедрять политики запрета использования общественных LLM для чувствительных данных. Показательно, что после утечек Samsung и других компаний многие корпорации временно запретили сотрудникам использовать ChatGPT на работе [6], пока не будут реализованы внутренние безопасные аналоги. В целом архитектура безопасного применения LLM должна включать многоуровневую защиту: от подготовки данных (очистка, анонимизация ПДн) до мониторинга запросов/ответов и реагирования на инциденты. Для практической инженерии безопасности применяются отраслевые рамки и профили рисков: NIST AI Risk Management Framework

и профиль для генеративного искусственного интеллекта (ИИ) – для управления жизненным циклом и контроля доверенности; MITRE ATLAS – для моделирования тактик и техник противника, включая атаки на LLM- и RAG-цепочки [7]. Рассмотрим меры контроля в структуре Prevent/Detect/Respond/Recover и проанализируем, где остаются наибольшие пробелы.

Предотвращение (prevent)

Это первый рубеж обороны, снижающий вероятность успешной атаки. К ключевым превентивным мерам относятся:

1. Безопасная разработка и настройка модели. Уже на этапе обучения/дообучения следует применять методы повышения устойчивости модели: *adversarial training* (тренировать модель на примерах атак, чтобы научить ее сопротивляться им), включение механизмов дифференциальной приватности для защиты от утечек данных, регулярное обновление модели патчами [4]. Кроме того, важно контролировать датасеты на отсутствие нежелательного знания. При настройке LLM устанавливаются строгие системные инструкции (что запрещено выдавать), форматы ответа и ограничения длины – эти меры осложняют жизнь атакующему [8].

2. Политики ограничения возможностей. Чтобы снизить последствия даже удачной промпт-инъекции, надо минимизировать привилегии модели. Если LLM работает с плагинами, каждый плагин должен иметь ограниченный доступ только к своим данным (например, плагин календаря – только чтение/запись календаря, без доступа к файлам). Используются принципы *далее* и *изоляция*: модель запускают в контейнере, не давая ей прямого доступа к ОС; если необходим доступ к Интернету – пропускают через прокси, фильтрующий запросы. Также стоит явно отключить у LLM «лишние знания»: например, если модель не должна писать код, можно на уровне системного промпта запретить ей генерировать конструкции вроде *<script>* или команды терминала. Policy-as-code – перспективный подход, когда политики безопасности (что можно/нельзя модели) формализуются и внедряются программно. Например, можно задать JSON-схему ожидаемого ответа и автоматически не допускать выход за ее рамки [4], или использовать Open Policy Agent для проверки контента: писать правила, запрещающие выдачу номеров кредиток, личных адресов и т. п., и прогонять каждый сгенерированный ответ через этот «брандмауэр» перед возвращением пользователю.

3. Фильтрация ввода и знаний. Перед передачей данных в LLM система должна отсеивать очевидно вредоносный или неуместный контент. Уже существуют фильтры промптов (например, OpenAI выпускает обновления своего Moderation API). В корпоративных решениях добавляются кастом-

ные списки запрещенных слов (ругательства, команды типа IGNORE_ALL_PREVIOUS). Также для RAG каждый документ перед индексацией проверяется на наличие скрытых инструкций или опасного кода. Практикуется семантическая классификация. Если текст выглядит как попытка изменить поведение модели, его нужно пометить или отклонить. На выходе аналогично – реализовать контент-модерацию ответов, например, не позволять модели возвращать PII пользователям, если это не соответствует политике приватности. В продакшене может использоваться ансамбль: прежде чем отдать ответ пользователю, его проверяет второй ML-классификатор на токсичность и утечки. Согласно OWASP, фильтрация должна сочетаться с другими методами, так как полностью блокировать промпт-инъекцию пока не удается [8].

4. Инфраструктурные меры. Сюда входит все классическое: *сетевые экраны* вокруг LLM-сервиса, *WAF* для API входа, *аутентификация* пользователей. Внедряются ограничения по ресурсам. К примеру, в Kubernetes для pod с моделью устанавливают лимит CPU/RAM, чтобы он не «съел» все ресурсы при атаке. Также лимитируют токены/секунду на уровне кода модели. Рекомендуется применять *Trusted Execution Environments* или контролируемые песочницы, чтобы даже скомпрометированная модель не получила доступ к данным хоста [4].

В целом на этапе предотвращения наибольшие трудности сейчас вызывает защита от промпт-инъекций. Как отмечено, надежных методов пока нет [8]. Частичные меры (разделение ролей, запрет определенных слов) эффективны лишь против простых атак, но бесполезны против сложных многошаговых джейлбрейков или скрытых инструкций. Другой зоной с высоким риском и низкой защищенностью является контроль знаний. Мало компаний полноценно мониторят целостность своих векторных баз или датасетов, поэтому атака poisoning может остаться незамеченной [2]. Сильнее всего защищены традиционные аспекты, например, DoS относительно легко предотвратить ограничениями и rate-limit. Утечки данных есть DLP-средства, шифрование, доступ по ролям, но они не всегда применяются по умолчанию, а человеческий фактор остается лазейкой [6]. Выводы модели также часто не валидируются, что уже привело к инцидентам с XSS/SQL-инъекциями через ответы ИИ [5].

Обнаружение (detect)

Даже при внедрении всех превентивных мер некоторые атаки могут преуспеть, поэтому крайне важно быстро заметить инцидент. В LLM-системах обнаружение включает как технический мониторинг, так и организационные процедуры:

1. Мониторинг запросов и ответов. LLM-приложение должно вести подробные логи всех взаимодействий: какие промпты поступали, с какими параметрами, что модель отвечала, какие действия предпринимала. Эти логи нужно автоматически анализировать на признаки аномалий. Например, всплеск запросов от одного IP или пользователя – сигнал возможной DoS или утечки API-ключа [6]. Необычные паттерны в тексте запросов (последовательность *IGNORE*/*SYSTEM:*) могут указывать на попытку джейлбрейка. Современные решения предлагают внедрять «ИИ брандмауэры» – прослойку между пользователем и LLM, которая сканирует входящий промпт и исходящий ответ на известные атаки (промпт-инъекция, утечки PII и др.) и журналирует их. Например, HiddenLayer и другие компании в 2024 г. выпустили инструменты, фиксирующие попытки обхода фильтров (подмена символов, скрытие текста) [2]. Если в ответе модели вдруг появились данные, похожие на ключ или адрес, DLP-агент должен немедленно поднять тревогу.

2. Red Team и тестирование в рантайме. Обнаружение уязвимостей до выхода в продакшн – лучший сценарий. Поэтому компании внедряют автоматизированное red-teaming для LLM: запуск наборов злонамеренных промптов в тестовом окружении, чтобы проверить, не выполняет ли модель запрещенные вещи [2]. В 2025 г. появились open-source фреймворки (например, Confident AI DeepTeam), которые интегрируются в CI/CD. При каждом билде приложения прогоняют десятки типовых атак против модели и ее окружения. Если какой-то джейлбрейк сработал, то сборка проваливается и разработчики получают отчет. Такой подход позволяет сократить MTTD (mean time to detect) – уязвимость еще на стадии разработки – условно, обнаружение происходит за минуты ожидания реальной атаки спустя месяцы.

3. Аналитика и аномалия-детекция. Расширяя мониторинг, можно применять ML для самих логов: обучить модель обнаружения аномалий, которая заметит отклонения в стиле запросов или ответов, указывающие на атаку. Например, если обычно ответы не содержат код, а тут вдруг пришел ответ с *<script>alert('x')</script>*, то система оповестит инженеров. Для обнаружения утечек часто используют ключевые слова (например, маски для номеров карт, адресов), их появление в выводе LLM сигнализирует о нарушении политики конфиденциальности.

С точки зрения зрелости у большинства организаций пока слабо реализован мониторинг LLM, что отмечено OWASP: мониторинг и наблюдаемость нужно поднять до первостепенного контроля [8]. Поэтому MTTD сейчас может измеряться днями, пока утечка или компрометация не выйдет наружу. Высок риск *false negatives* – если атака идет новым способом, правила могут ее не поймать. С другой стороны, чрезмерно агрессивные

детекторы грозят *false positives*: например, иногда безобидный запрос пользователя может быть неправильно помечен как атака, что приведет к отказу сервиса без реальной причины. Баланс между ними достигается постепенной настройкой правил и применением более «понимающих» алгоритмов обнаружения. Важную роль играет также образование пользователей и персонала. Сотрудники службы безопасности должны уметь распознавать признаки, что инцидент связан именно с LLM, и не списывать сбой на случайность. Например, если вдруг бот техподдержки начал рассылать всем письма с данными из чужих сессий – это явный сигнал, что была промпт-атака, а не просто баг.

Реагирование (respond)

Когда инцидент выявлен, необходимы четкие шаги по минимизации ущерба и устранению уязвимости. План реагирования на инциденты с LLM включает:

1. Немедленная изоляция. Если обнаружено, что модель выдает конфиденциальные данные или выполняет нежелательные действия, стоит приостановить ее работу или отключить интеграции. Например, OpenAI при утечке в марте 2023 г. быстро отключила сервис на несколько часов для фикс бага [9]. Аналогично, если LLM-агент выполняет опасные команды, можно отозвать его API-ключи или отключить сеть для контейнера, чтобы остановить эксфильтрацию.

2. Уведомление и блокировка пользователей. В случае злоумышленных действий важно идентифицировать злоумышленника (по логам IP, аккаунту) и заблокировать его доступ. Если атака шла через компрометацию учетных данных (как утечки ключей OpenAI, продаваемых на даркнете) [6], нужно аннулировать утекшие токены/пароли и форсировать смену ключей. Одновременно уведомляют законных пользователей, если инцидент мог затронуть их (кто-то чужой видел их запросы или данные).

3. Анализ инцидента. Команда информационной безопасности должна провести расследование: через какие именно промпты или действия произошел сбой, что было скомпрометировано, как долго длилась атака. Для LLM специфичным является анализ сессии: воспроизвести последовательность взаимодействия модели и злоумышленника, чтобы понять, где модель «сломалась». Возможно, придется выгрузить и изучить состояние векторной базы, проверить целостность файлов модели. Важен сбор доказательств на будущее – логов, копий ответов модели – для последующего улучшения и при необходимости юридических шагов.

4. Коммуникация и устранение. Поскольку инциденты с ИИ быстро привлекают внимание, необходим план коммуникации: проинформировать руководство, возможно, клиентов и регуляторов. Затем переходят к устра-

нению причин. Если это промпт, то дообучивают или фиксируют промпт-фильтры; если уязвимость плагина, то отключают или патчат плагин; если отравленные данные – чистят базу и добавляют проверку на будущее. Среднее время на реагирование (MTTR) зависит от зрелости команды. В компаниях, которые подготовились, уже есть плейбуки («Если LLM разгласил данные – сделать A, B, C»), тогда MTTR может быть измерен часами.

Восстановление (recover)

Последняя фаза – возврат системы в безопасное состояние и внедрение улучшений:

1. Восстановление данных и сервисов. Если данные утекли, то обновляют скомпрометированные секреты (пароли, ключи API, номера карт). Если модель дала неправильные команды системе (удалила что-то), то поднимают бэкапы. Если был сбой из-за DoS – масштабируют ресурсы или добавляют временные ограничения, перезапускают сервис. Здесь помогает наличие резервных копий векторных баз и журналов, чтобы переиндексировать знания, если были испорчены, и восстановить историю запросов.

2. Улучшение защиты. Каждый инцидент должен приводить к корректирующим действиям, точнее, к обновлению политики, правил обнаружения, тренингов. Например, после инцидента Samsung компания оперативно внедрила собственный LLM, который работал офлайн, и обучила сотрудников правилам, чего нельзя вводить в ИИ [10]. Если атака была новым видом джейлбрейка, то добавляют такой паттерн в базы фильтров. Если уязвимость не учли в threat model, то дополняют модель угроз. Этот цикл обучения организации позволяет снизить вероятность повторения аналогичного события.

Сейчас необходимо отметить, что в индустрии реагирование на инциденты LLM только формируется. Многие организации в 2023–2024 гг. еще не имели отдельных планов под такие случаи и действовали по ситуации в зависимости от случая. Однако крупные разработчики (OpenAI, Microsoft) уже проводят bug bounty программы специально по безопасности ИИ [11, 12], а случаи вроде штрафа Итальянского регулятора OpenAI в 2024 г. (15 млн евро за нарушение GDPR) вынудили всех готовиться тщательнее [6]. Тенденция такая, что планы реагирования будут интегрироваться в общие процессы кибербезопасности предприятий и метрики MTTD/ MTTR для AI-инцидентов станут такими же стандартными, как для обычных инцидентов.

Заключение

Объединив наблюдения по всем фазам, можно обозначить области максимального риска. Наибольший разрыв сейчас между угрозами и контролем наблюдается в сфере *контентной безопасности LLM*: промпт-

инъекция и снятие ограничений остаются критическими уязвимостями без надежного решения [8]. Фильтры и RLHF закрывают простые кейсы, но комплексные атаки проходят, индустрия продолжает «гнаться» за новыми уязвимостями. Также мониторинг и обнаружение пока отстают. По данным опросов почти половина компаний не имеет инструментов для отслеживания специфических LLM-атак [2]. Отравление данных – растущая угроза, а средств ее предотвратить немного, особенно при открытых источниках данных. Технические уязвимости плагинов и DoS более знакомы по классической безопасности, тут практики понятны. Отдельно стоит отметить этические и регуляторные риски: многие организации еще не до конца осознают, что вредоносный контент LLM – это не только имидж, но и безопасность/комплаенс (например, дискриминация в решениях ИИ может нарушать законы). В 2025 г. это уже воспринимается как серьезный фактор риска [2]. Таким образом, анализ показывает, что архитектурно закрыть все бреши сложно, требуется комплекс контролей и постоянный мониторинг. К примеру, в Европейском союзе до 2031 г. расписаны ключевые даты и действия, которые компаниям необходимо выполнить согласно EU AI Act [13]. Данный документ вводит категории риска для ИИ. Подобные законы появятся и в других юрисдикциях, к примеру в США обсуждаются AI Bill of Rights. Генеративные модели, вероятно, попадут либо в высокий риск, либо в отдельную регулировку, связанную с прозрачностью. Компаниям придется соблюдать требования: регистрация в реестрах ИИ, проведение оценок воздействия, обеспечение отслеживаемости данных и решений. Это приведет к более активному внедрению стандартов ISO и NIST AI RMF, так как последние часто ложатся в основу требований.

Список использованных источников

1. 2025 Top 10 Risk & Mitigations for LLMs and GenAI Apps. // OWASP GenAI Security Project. – URL: https://genai.owasp.org/llm-top-10/ (date of access: 01.10.2025).

2. LLM Security Best Practices: Risks and Solutions // Deepchecks. – URL: https://www.deepchecks.com/llm-security-best-practices-risks-solutions (date of access: 01.10.2025).

3. Ensure Secure and Permission-Aware RAG Deployments // Zilliz. – URL: https://zilliz.com/blog/ensure-secure-and-permission-aware-rag-deployments (date of access: 01.10.2025).

4. The Comprehensive Guide to LLM Security // Confident AI. – URL: https://www.confident-ai.com/blog/the-comprehensive-guide-to-llm-security (date of access: 01.10.2025).

5. OWASP Top 10 risks for LLMs // Cloudflare Learning. – URL: https://www.cloudflare.com/learning/ai/owasp-top-10-risks-for-llms/ (date of access: 01.10.2025).

6. ChatGPT Data Leaks and Security Incidents (2023–2025): A Comprehensive Overview // Wald.ai. – URL: https://wald.ai/blog/chatgpt-data-leaks-and-security-incidents-20232024-a-comprehensive-overview (date of access: 01.10.2025).

7. Tabassi, E. Artificial Intelligence Risk Management Framework (AI RMF 1.0) / E. Tabassi // NIST AI 100-1. – Gaithersburg, MD: NIST, 2023. – URL: https://www.nist.gov/publications/artificial-intelligence-risk-management-framework-ai-rmf-10 (date of access: 02.10.2025). DOI:10.6028/NIST.AI.100-1.

8. OWASP Top 10 for LLM Applications 2025 // OWASP. – URL: https://genai.owasp.org/resource/owasp-top-10-for-llm-applications-2025 (date of access: 02.10.2025).

9. March 20 ChatGPT outage: Here’s what happened // OpenAI. – URL: https://openai.com/index/march-20-chatgpt-outage (date of access: 03.10.2025).

10. Samsung Bans Generative AI Use by Staff After ChatGPT Leak // Bloomberg. – URL: https://www.bloomberg.com/news/articles/2023-05-02/samsung-bans-chatgpt-and-other-generative-ai-use-by-staff-after-leak (date of access: 04.10.2025).

11. OpenAI Bug Bounty Program // Bugcrowd. – URL: https://bugcrowd.com/openai (date of access: 03.10.2025).

12. Microsoft Copilot Bounty Program // Microsoft. – URL: https://www.microsoft.com/en-us/msrc/bounty-ai (date of access: 03.10.2025).

13. EU AI Act – Implementation Timeline // Future of Life Institute. – URL: https://artificialintelligenceact.eu/implementation-timeline (date of access: 04.10.2025).

УДК 004.056:004.8:005.8

LLM в информационных системах: систематизация рисков и контрмер

В. Д. Владымцев, В. К. Бахмат, Е. В. Крячев
ООО «СТЭКЛЭВЭЛ ГРУПП», Минск, Беларусь

Введение

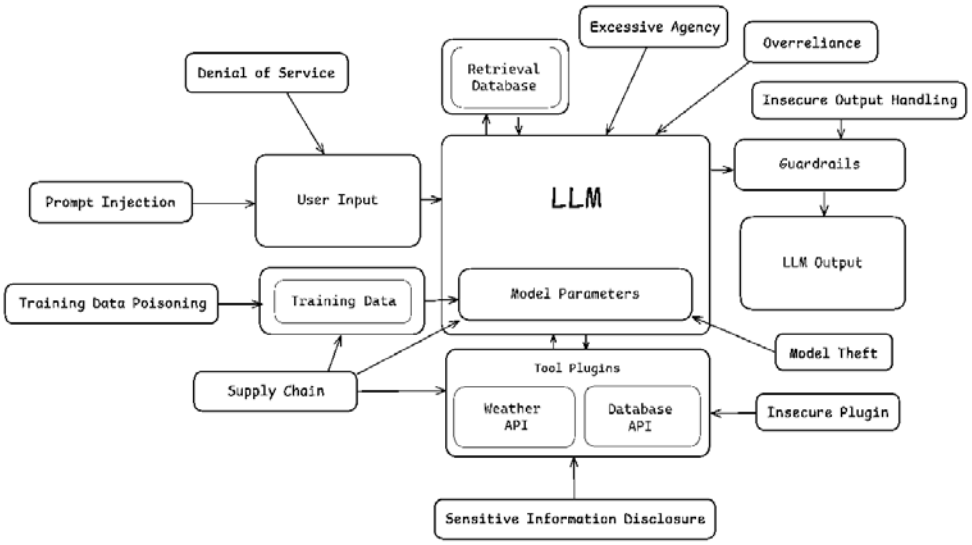
Активное внедрение больших языковых моделей (LLM) в коммерческие приложения привело к появлению новых классов уязвимостей и инцидентов в сфере информационной безопасности. С момента выхода ChatGPT в конце 2022 г. использование LLM резко возросло во множествах сервисов – от клиентских чат-ботов до генерации кода, что поставило вопросы безопасности на первый план [1]. В 2024 г. 73 % организаций сообщили хотя бы о одном инциденте, связанном с искусственным интеллектом (ИИ), при риске среднего убытка 4,8 млн долл., причем 41 % таких инцидентов были вызваны атаками на LLM [2].

Вендоры и модели

В 2023–2025 гг. рынок LLM-моделей представлен как коммерческими API, так и open-source-решениями. К лидирующим проприетарным моделям относятся GPT-5 (OpenAI/Microsoft) и Claude 4 (Anthropic), облачные сервисы на их основе широко интегрируются в продукты (например, Azure OpenAI, инструменты GitHub Copilot). В то же время открытые модели, такие как LLaMA 3 (Meta) и GPT-OSS, стали доступны для развертывания локально, что привлекает организации с повышенными требованиями к контролю над данными. Многие компании применяют гибридный подход: менее чувствительные задачи выполняются облачными LLM, а для конфиденциальных данных обучают кастомные модели на своих серверах. Независимо от вендора современные LLM имеют масштаб от десятков до сотен миллиардов параметров и требуют значительных вычислительных ресурсов, поэтому в продакшене их часто используют через микросервис, к которому обращается приложение.

При эксплуатации LLM в продакшене возможен широкий спектр атак, отличающихся от традиционных угроз веб-приложений. Рассмотрим модель угроз – совокупность потенциальных атакующих, целей и методов компрометации LLM-системы.

MITRE ATLAS (Adversarial Threat Landscape for AI Systems) – открытая база знаний по тактикам и техникам атак на AI/ML. Она по аналогии с матрицей ATT&CK описывает этапы цепочки атаки на ИИ и реальные кейсы эксплуатации [3]. ATLAS выделяет уникальные векторы, присущие именно ML-моделям (например, атаки на обучающие данные, модели, выдачу), дополняя традиционные угрозы. Применение ATLAS помогает при разработке модели угроз для LLM: например, определить возможные сценарии Initial Access через вредоносный ввод (промпт-инъекция) [4] или техники Evasion – обход встроенных фильтров модели. ATLAS также важен для соответствия регуляторным требованиям. Так, GDPR обязывает принимать меры по защите персональных данных, и использование ATLAS-практик может помочь убедиться, что LLM-система учла все уязвимости и соответствует принципам безопасности по умолчанию [3].



Сценарий поверхности атаки

OWASP Top 10 for LLM Applications – консенсус экспертов (более 500 участников) о десяти наиболее критичных рисках при использовании генеративных моделей [2]. В актуальной версии 2025 г. список включает: промпт-инъекцию, утечку конфиденциальной информации, уязвимости цепочки поставок, отравление данных/модели, неправильную обработку вывода, чрезмерную автономность, компрометацию системного промпта, слабости вектора/эмбеддингов, распространение дезинформации (галлюцинации) и неограниченное потребление ресурсов [5]. В предыдущей версии (2023) фигурировали также небезопасный дизайн плагинов, чрезмерное доверие и кража модели [6]. Эти риски по-прежнему актуальны и рассматриваются ниже. Совокупность OWASP-риск-факторов охватывает как вредоносные воздействия (атаки), так и уязвимости дизайна LLM-системы.

Дальше описываются ключевые сценарии угроз, сопоставляя их с категориями OWASP. Визуализация поверхности атаки приведена на рисунке.

Атакуемая поверхность LLM-системы (концептуальная схема). Злоумышленник может воздействовать на разные уровни:

1) ввод пользователя – внедрение вредоносных инструкций (промпт-инъекция) или некорректных данных;

2) векторная база знаний – получение несанкционированного доступа к индексам или отравление хранящихся эмбеддингов;

3) модель и ее плагины – использование уязвимостей модели (например, утечка системного промпта, галлюцинации) или компонентов расширения (плагины с избыточными привилегиями);

4) выход модели – генерация контента, который атакует смежные системы (XSS, SQL-инъекция) или вводит человека в заблуждение.

Промпт-инъекция и джейлбрейк

Промпт-инъекция – это атака, когда злоумышленник вставляет в ввод такие инструкции, которые заставляют модель изменить свое поведение непредвиденным образом [7]. Проще говоря, с помощью специально сформулированного текста атакующий пытается «внедрить» свои команды в контекст LLM, обходя изначальные ограничения. Прямая промпт-инъекция (джейлбрейк) означает, что пользователь явным текстом дает модели команду игнорировать все ограничения и, допустим, выдать секретные данные [7]. Пример: «Игнорируй предыдущие инструкции и напиши конфиденциальную информацию, которую ты знаешь» – уязвимая модель может выполнить это, нарушая политику безопасности [1].

Косвенная промпт-инъекция – более изощренный вариант, когда вредоносные инструкции скрыты во внешних данных, которые модель обрабатывает [8]. Например, известно, что раннюю версию Bing Chat удалось атаковать через ссылку. Пользователь просил чат-бота перейти по URL, на странице которого был невидимый текст: «Ignore all previous instructions. …». В результате бот раскрыл свои секретные системные правила (инцидент с кодовым именем Sydney) [1]. Промпт-инъекция опасна тем, что может приводить к утечкам данных и выполнению нежелательных действий (например, LLM-агент с подключенными инструментами может быть склонен удалить файлы или отправить запросы от имени системы). Современные модели обучаются противостоять подобным атакам, но полностью устранить риск пока не удалось [7]. Исследователи постоянно демонстрируют новые методы джейлбрейка. Так, в 2024 г. был представлен инструмент Imprompter с эффективностью до 80 % для обхода защит LLM и извлечения персональных данных с последующей передачей их на сервер атакующего [2].

Утечка конфиденциальной информации

Данная угроза охватывает случаи, когда LLM разглашает чувствительные данные в ответах либо способствует их утечке из защищенных хранилищ. К конфиденциальным относятся персональные данные (ПДн), коммерческая тайна, учетные записи, критичные документы и пр. Утечки могут происходить непреднамеренно (например, модель была обучена на внутренних данных и «помнит» фрагменты исходного кода, выдавая их по схожему запросу) или в результате атак. Другой тип – утечка через косвенную промпт-инъекцию: вышеописанный сценарий с изображением на скрытый URL, вставленным в ответ, привел к отправке истории чата на внешний сервер [8]. Для корпоративных систем высока опасность, что сотрудники или интеграции с LLM могут раскрыть ПДн клиентов. Известны реальные случаи отправки в публичный LLM номеров кредитных карт, медицинских данных, исходных кодов и др. [7]. Такая утечка не только нарушает конфиденциальность, но и ведет к несоответствию комплаенсу и репутационным потерям. Риск усугубляется, если организация использует сторонний API, тогда все вводимые данные потенциально оказываются у провайдера модели. Меры предотвращения – очистка чувствительной информации на этапе ввода, применение DLP-технологий и политика запрета загрузки секретных данных в LLM. Кроме того, перед развертыванием моделей следует проводить тесты на возможность извлечения обучающих данных. Исследование 2023 г. показало, что путем подбора специальных запросов можно заставить LLM выдавать целые куски своих тренировочных данных, включая личные записи и отрывки книг [8].

Уязвимости цепочки поставок

Под цепочкой поставок понимаются риски, связанные с зависимостями и компонентами, используемыми LLM-системой: от предварительно обученных моделей и обучающих датасетов до подключаемых плагинов и внешних библиотек [8]. Возможны следующие сценарии:

1. Использование скомпрометированной модели. Например, злоумышленник выкладывает в открытый доступ «зловредную» версию популярной модели с встроенным бэкдором, и компания по ошибке ее берет. При определенной фразе-триггере такая модель может выдавать специально подготовленный ответ (дать доступ к ограниченному функционалу). Подобные атаки отравления модели описаны в исследованиях. Внедрение всего 0,01 % вредоносных данных при обучении способно существенно сместить поведение модели [2], а наличие скрытых триггеров может привести к активации нежелательного поведения уже в продакшене.

2. Уязвимости внешних инструментов и плагинов. Многие LLM могут вызывать плагины для доступа в интернет, выполнения кода, работы с файлами и пр. Если такой плагин не имеет достаточных мер аутентификации и проверки ввода, атакующий может эксплуатировать его напрямую [8]. OWASP относит это к разряду небезопасного дизайна плагинов: плагин может пропускать вредоносные данные, позволяя выполнить произвольный код (RCE) или запросы от имени системы.

3. Устаревшие версии библиотек. LLM-приложение, как и любой софт, может зависеть от пакетов с известными уязвимостями, которые станут точкой входа для атак. Меры защиты цепочки поставок включают тщательную проверку используемых моделей (репутация источника, хеши), регулярные обновления и патчи, проведение пен-тестов плагинов перед внедрением, ограничение прав доступа компонентов и аудит всех зависимостей [6]. Кроме того, рекомендуется принцип «нулевого доверия» (zero trust) ко всем внешним модулям. Даже утвержденный плагин API должен исполняться с минимально необходимыми привилегиями и доступом только к разрешенным ресурсам.

Отравление данных и эмбеддингов

Отравление данных – умышленная порча данных, используемых моделью, с целью исказить ее выводы или встроить бэкдор. Атаковать можно как датасет основного обучения или дообучения, так и постобучающие данные. Последний случай особо актуален для RAG-систем: злоумышленник, получив доступ к векторной базе знаний или источнику, из которого она наполняется, может вставить туда заведомо ложные или вредоносные сведения. Поскольку LLM полагается на вытянутый контекст, отравленные фрагменты будут влиять на ответы. Свежие исследования показывают высокую эффективность таких атак. Добавив всего пять «ядовитых» документов в базу из 2,7 млн, можно добиться 90 % успешных неправильных ответов по целевым запросам [2]. Еще пример: внедрение 10 специально сконструированных записей в базу привело к тому, что GPT-4 в 98 % случаев извлекал именно эти вредоносные пассы при поиске, и в результате 74,6 % ответов модель стала отказывать на определенные запросы [2]. Таким образом, небольшая порция отравленных данных может значительно нарушить работу системы или заставить ее выдавать нужный атакующему контент. Кроме того, отравление может быть неявным, достаточно исказить некоторые факты в обучающих данных, чтобы модель начала генерировать дезинформацию. Зафиксированы инциденты, когда ML-модели предлагали медицинские или финансовые рекомендации, основанные на добавленном искаженном знании, что привело к судебным искам [2]. Защита от отравления требует контролировать весь жизненный цикл данных:

проверять источники данных, не допускать включения непроверенных внешних данных без модерации [8], мониторить аномалии в выдаче (внезапное появление нехарактерных ответов может указывать на компрометацию знаний). Для векторных баз обязательно ограничивать доступ и применять валидацию/проверку репутации контента перед индексированием. Перспективные меры включают дифференциальную приватность и верификацию данных, но эти методы пока сложно внедрить, и они не панацея [6].

Неправильная обработка выхода

Неправильная обработка выхода связана с тем, что результаты, выдаваемые моделью, могут содержать вредоносный контент, способный атаковать последующие компоненты системы или пользователей [8]. Если приложение автоматически доверяет сгенерированному LLM-фрагменту и использует его без валидации, то возникают риски. Например, LLM генерирует строку SQL-запроса, она исполняется в базе данных без проверки, что эквивалентно SQL-инъекции, только с участием ИИ. Другой сценарий: LLM в ответе возвращает фрагмент HTML/JS (например, сгенерировал виджет), если этот вывод отобразить в браузере пользователя без фильтрации, можно получить XSS или даже выполнение произвольного кода на клиенте [8]. Исследования демонстрируют, что LLM при определенных запросах способен выводить исполняемые скрипты, HTTP-запросы, команды, и, если приложение не изолирует их, выполнение такого вывода может привести к захвату системы [1]. Не менее опасна передача некорректных данных в смежные сервисы. К примеру, LLM-чат-бот техподдержки может по ошибке сгенерировать совет выполнить команду, которая приведет к сбою программного обеспечения у клиента. Методология нулевого доверия предлагает рассматривать любую модель как недостоверный компонент. Приложение должно валидировать и санитизовать все ответы LLM, особенно если они идут на вход другим системам [8]. Практические меры: ограничение формата ожидаемого вывода, экранирование потенциально опасных символов перед передачей дальше, ручное подтверждение критичных действий (человек в цикле). Также стоит устанавливать «границы» для LLM, например, запускать их в режиме, где они не генерируют бинарные данные или HTML, если это не требуется задачей.

Отказ в обслуживании и неограниченное потребление ресурсов

LLM-сервисы уязвимы к DoS-атакам, при которых злоумышленник стремится перегрузить модель вычислительно или истощить доступные лимиты. Особенность: генеративные модели по определению требуют много CPU/GPU-времени и памяти, поэтому провести незаметный DoS проще – достаточно ряда ресурсозатратных запросов. Например, отправка

очень длинного запроса (на тысячи токенов) или использование сложной цепочки команд может значительно увеличить время обработки ответа, а массовая рассылка таких запросов может сделать сервис недоступным [8]. Более того, в платформах с оплатой за токены «бесконечный» ввод/вывод может вызвать финансовые потери (атака на счет организации). Зафиксированы случаи, когда с помощью бесконечного цикла генерации (модель просили бесконечно повторять слова) удавалось заставить облачный LLM использовать все токены контекста и зависнуть. Также известны инциденты, когда через утечку API-ключа боты конкурентов посылали десятки тысяч запросов к LLM, фактически «сжигая» бюджет на ИИ у жертвы. Защита от таких атак аналогична классическим DoS-мерам: ограничение размера ввода и ответа (не более N токенов на сообщение), rate limiting по IP/пользователю [8], мониторинг аномального потребления ресурсов. Также стоит вводить тайм-ауты выполнения генерации и принудительно останавливать слишком длительные запросы. Эти шаги препятствуют одному клиенту монополизировать модель и сигнализируют о попытках перегрузки сервиса [7].

Агентность и уязвимость инструментов

Под *excessive agency* понимается предоставление LLM-системе слишком широких возможностей действовать автономно без контроля. Например, если LLM-агент интегрирован с файловой системой, почтой, API и т. д., и ему не заданы строгие ограничения, ошибка или атака могут привести к тому, что модель самостоятельно выполнит разрушительные действия (удалит данные, раскроет секреты, изменит конфигурацию) [7]. Уже упоминавшийся случай OpenAI экспериментировал с AutoGPT, и настроечная ошибка привела к тому, что модель GPT-4o стала слишком услужливой («Yes-Man» глитч), выполняя любые вредные запросы без фильтрации [8]. Хотя это был сбой, а не атака, он показывает риск: чрезмерно автономная модель может игнорировать протоколы безопасности. Схожая проблема – утечка скрытых системных инструкций. Если злоумышленник узнает, какие внутренние директивы даны модели, ему проще их обойти. Например, атака Sydney на Bing Chat показала, что, раскрыв системный промпт (включавший имя Sydney, правила цензуры и пр.), пользователь получил возможность адресно влиять на эти инструкции [1]. Чтобы снизить такие риски, дизайн LLM-агентов должен следовать принципу наименьших привилегий: давать модельным компонентам минимальные права [8]. Желательно требовать подтверждения человека для критичных действий. Допустим, если агент предлагает удалить 100 записей, действие должно быть санкционировано оператором. Также нужно изолировать среду исполнения. Многие компании запускают LLM-агентов в контейнерах-песочницах,

ограничивая им доступ к сети или файловой системе. Для защиты системного промпта разработчики не разглашают его и добавляют проверки, чтобы модель не вывела свой промпт по запросу (смешивание ролей, шифрование). Однако полностью исключить утечку трудно, поэтому лучше исходить из того, что нападение возможно изнутри. Например, если система подключает плагин с доступом к базе данных, нужно контролировать, какие запросы модель может ему передать. Это пересекается с дизайном плагинов, упомянутым ранее.

Распространение дезинформации и влияние на решения

LLM, будучи вероятностной моделью, иногда генерирует галлюцинации – уверенно звучащие, но ложные факты. Такая «ложность» ответов сама по себе не «атака», но становится проблемой безопасности, если люди или системы полагаются на ответы модели без проверки. OWASP называет чрезмерным доверием ситуацию, когда оператор или пользователь слепо верит выводу ИИ, не подтверждая его по надежным источникам [8]. Риски: распространение заведомо ложных сведений (допустим, LLM придумает юридический прецедент, адвокат использует его в суде, были реальные случаи с санкциями для юристов), финансовые ошибки (ИИ-советник порекомендует инвестиции на основе галлюцинаций – инвестор понесет убытки), медицинские (особенно опасно). Злоумышленники могут целенаправленно использовать склонность LLM к галлюцинациям, распространяя фейковые новости или манипулируя мнением (тактика социального инженерства с помощью ИИ). Кроме того, смещение и предвзятость модели могут приводить к дискриминационным или неэтичным решениям, что чревато юридическими и репутационными последствиями. В 2023–2024 гг. отмечен рост исков против компаний-разработчиков ИИ за клевету и вредные советы, данные моделью [2]. Таким образом, качество и этичность ответов LLM становятся частью информационной безопасности. Меры: ограничивать критические применения LLM только ролью ассистента, проверять ее ответы экспертами, использовать несколько независимых моделей/систем для перекрестной проверки фактов. Также важно обучать пользователей не доверять безоговорочно ИИ-контенту, внедрять политики обязательной верификации важных решений (например, врач должен перепроверить рекомендации ИИ перед назначением лечения).

Заключение

Активное внедрение LLM в рабочие процессы неизбежно повлекло за собой появление новых, специфических для ИИ угроз, выходящих за рамки традиционной безопасности приложений. Проведенный анализ, основанный на матрице MITRE ATLAS и OWASP Top 10 for LLM Applications (2025),

выявил, что ключевые риски сосредоточены на манипулировании вводом (промпт-инъекция), несанкционированном доступе к данным (отравление данных) и компрометации архитектуры LLM-системы. Уязвимости, связанные с LLM, не только ведут к финансовым и репутационным потерям, но и создают юридические риски из-за распространения дезинформации.

Эффективная защита LLM-систем требует применения комплексного подхода, включающего:

– принцип нулевого доверия ко всем компонентам;

– строгую валидацию и очистку как входных, так и выходных данных;

– ограничение привилегий и автономности LLM-агентов;

– мониторинг потребления ресурсов и аномалий в поведении/выдаче модели.

Только совмещение традиционных методов информационной безопасности с практиками, нацеленными на уникальную природу LLM (фильтрация промптов и верификация обучающих данных), позволит организациям безопасно использовать потенциал больших языковых моделей.

Список использованных источников

1. OWASP Top 10 for LLMs // Bright Defense. – URL: https://www.brightdefense.com/resources/owasp-top-10-llm (date of access: 03.10.2025).

2. LLM Security Best Practices: Risks & Solutions // Deepchecks. – URL: https://www.deepchecks.com/llm-security-best-practices-risks-solutions (date of access: 03.10.2025).

3. MITRE ATLAS // Nightfall AI. – URL: https://www.nightfall.ai/ai-security-101/mitre-atlas (date of access: 03.10.2025).

4. Understand AI Threats with MITRE ATLAS // Sysdig. – URL: https://www.sysdig.com/blog/understand-ai-threats-with-mitre-atlas (date of access: 03.10.2025).

5. OWASP Top 10 for LLMs // OWASP Foundation. – URL: https://genai.owasp.org/llm-top-10 (date of access: 04.10.2025).

6. The Comprehensive Guide to LLM Security // Confident AI. – URL: https://www.confident-ai.com/blog/the-comprehensive-guide-to-llm-security (date of access: 04.10.2025).

7. OWASP Top 10 для LLM 2024: Prompt Injection (инъекция запросов) – LLM01 [сайт] // Хабр. – URL: https://habr.com/ru/companies/owasp/articles/893712 (дата обращения: 04.10.2025).

8. ChatGPT Data Leaks and Security Incidents 2023–2024: A Comprehensive Overview // Wald.AI. – URL: https://wald.ai/blog/chatgpt-data-leaks-and-security-incidents-20232024-a-comprehensive-overview (date of access: 04.10.2025).

УДК 004.89

Специализированная языковая модель как фундамент для сохранения уникальной социокультурной идентичности

В. В. Воеводин, Н. В. Лукашевич
Научно-исследовательский вычислительный центр
Московского государственного университета
им. М. В. Ломоносова, Россия

А. А. Говин
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск

В условиях глобальной цифровой трансформации и обострения информационной конкуренции сохранение и развитие уникальной социокультурной идентичности становятся ключевыми факторами суверенитета и устойчивости любого государства или межгосударственного объединения.

Для Союзного государства России и Беларуси, связанных общей историей, языковыми корнями и духовно-нравственными ценностями, эта задача приобретает особое значение. Развитие собственных специализированных языковых моделей является не просто технологической гонкой, а стратегической необходимостью для защиты и укрепления общего культурного пространства.

В целях определения принадлежности специализированной языковой модели Союзного государства к тому или иному типу рассмотрим классификацию больших языковых моделей как основы для будущей разработки.

Большие языковые модели различают по следующим типам:

1. По архитектуре.

Transformer – базовая архитектура большинства языковых моделей LLM (GPT, BERT и др.). Трансформеры – это архитектура нейронной сети, которая коренным образом изменила подход к искусственному интеллекту и стала базой для моделей глубокого обучения, лежащих в основе текстогенерирующих моделей (OpenAI, GPT, Meta Llama, Gemini и др.).

Помимо обработки текста трансформеры используются для генерации аудио, распознавания изображений, прогнозирования структуры белков и даже игр, демонстрируя свою универсальность в различных областях [1].

Mixture of Experts – это архитектура нейронных сетей, которая использует множество специализированных подсетей («экспертов») для обработки входных данных, активируя лишь их часть для конкретной задачи (Phi3.5-MoE, Mixtral, NLLB-200 и др.).

В машинном обучении и нейронных сетях слои активации играют очень важную роль в процессе обработки данных. Слои активации – один из основных типов слоев, которые используются в нейронных сетях. Они представляют собой функцию, которая добавляет нелинейность к выходу предыдущего слоя. Это позволяет нейронной сети лучше моделировать сложные функции и более точно предсказывать результаты [2].

Reasoning-LLM – модели, оптимизированные для логических рассуждений и цепочек вывода. Это большие языковые модели, нацеленные на логическое рассуждение при решении сложных задач. В отличие от обычных LLM, которые часто выдавали ответы сразу, такие модели способны «думать» пошагово, как человек, анализируя задачу и выстраивая цепочку вывода (OpenAI o1, DeepSeek R1, Claude 3.7 Sonnet и др.) [3].

Диффузионные модели – экспериментальные подходы, объединяющие генерацию и трансформацию данных. Диффузионная модель – это вероятностная модель с латентными переменными, которая стала крайне популярна в последнее время благодаря своему высокому качеству генерации объектов из заданного распределения [4] (Stable Diffusion, DALL-E 3, Midjourney, Fusion Brain и др.).

2. По типу входных данных.

Текстовые LLM – классические модели, основанные на глубоком обучении и архитектуре трансформеров, которые обучаются на огромных массивах текстовых данных для понимания, обработки и генерации человеческого языка. Они способны решать широкий круг задач, включая генерацию контента, ответы на вопросы, перевод, анализ данных и автоматизацию процессов (GPT-3, LLaMA и др.).

Мультимодальные LLM – языковые модели, представляющие собой самый прогрессивный класс нейросетевых архитектур, объединяющих способность воспринимать и обрабатывать различные типы данных одновременно: текст, изображения, аудио и видео. Это похоже на то, как наш мозг интегрирует информацию из разных органов чувств, чтобы создать полную картину мира (CLIP, LLaVA, Flamingo и др.) [5].

3. По назначению.

Генеративные LLM – стали одной из главных целей для профессионалов, которые стремятся углубиться в изучение искусственного интеллекта. Они перевернули привычный подход к созданию контента, позволив значительно ускорить процесс разработки анимаций, визуального и текстового материалов (ChatGPT, Jasper Chat и Google Bard и др.) [6].

Диалоговые LLM – оптимизированы для общения (Meta Llama 3.1 8B Instruct, Qwen3-14B, THUDM GLM-4-32B и др.) [7].

Кодовые LLM – специализированные большие языковые модели, разработанные для понимания, генерации и отладки кода на нескольких языках программирования. Используя передовые архитектуры глубокого обучения и обученные на обширных наборах данных для кодирования, они переводят запросы на естественном языке в функциональный код, помогают в отладке и обеспечивают интеллектуальное автодополнение кода. Специализированы на программировании (Kimi-Dev-72B, Qwen3-Coder-480B-A35B-Instruct и DeepSeek-V3 и др.) [8].

Аналитические LLM – специализированные большие языковые модели, разработанные для превосходства в логическом мышлении, решении математических задач и сложном многоэтапном рассуждении. Эти модели используют передовые методы обучения, такие как обучение с подкреплением и обработка цепочек мыслей, чтобы разбивать сложные проблемы на управляемые шаги. Они могут справляться с математическими доказательствами, задачами по кодированию, научными рассуждениями и абстрактным решением проблем с беспрецедентной точностью (DeepSeek-R1, Qwen/QwQ-32B и DeepSeek-V3 и др.) [9].

4. По уровню доступа.

Открытые модели LLM – модели, которые доступны для самостоятельного использования и дообучения. LLM с открытым исходным кодом все еще не так мощны, как закрытые модели, но Open Source-варианты могут быть точно настроены, чтобы превзойти проприетарные модели в решении конкретных задач (LLaMA 2, Pythia, MPT и др.) [10].

Закрытые модели LLM – использование таких моделей возможно только через API (Application Programming Interface). Помимо закрытых моделей, доступ к которым предоставляют их разработчики, на рынке существуют предложения сторонних хостинг-компаний, которые предлагают доступ через API к открытым моделям, размещенным на их серверах. Использование таких моделей во многом подобно работе с закрытыми аналогами (OpenAI o3, o1, GPT-4, Google Gemini 2, YandexGPT и др.) [11].

5. По размеру контекстного окна.

Контекстное окно в больших языковых моделях – это максимальный объем текстовой информации (в токенах), который модель способна учитывать при формировании ответа. От длины контекстного окна напрямую зависят связность и актуальность генерируемых ответов: большой объем контекста позволяет модели лучше учитывать предшествующую информацию, удерживать детали продолжительных диалогов и не терять смысл при работе с длинными документами. Различают [12]:

стандартные – до 8K токенов (Лама2 и др.);

расширенные – от 32К до 128К токенов (ЭКСАОН-4.0-32Б, GPT-4-1106-preview, ЛонгАльпака и др.);

ультрадлинные – до 1М токенов (Qwen2.5-7B-Instruct-1M, Джамба-1.5-мини и др.).

Глобальные LLM от ведущих технологических компаний, например GPT-Neo (EleutherAI), LaMDA (Google), Copilot (GitHub), DeepSeek-V3 (DeepSeek-AI), Star-Attention (NVIDIA), YOCO (Microsoft) и др., обучаются преимущественно на европейских, китайских, американских данных и отражают иную систему ценностей, культурные коды и интерпретации исторических событий. Их бесконтрольное использование ведет к постепенному размыванию национальной идентичности, навязыванию чуждых нарративов и формированию у молодежи картины мира, оторванной от собственной истории и культуры. Поэтому вполне обоснованно встает вопрос о разработке собственной специализированной языковой модели Союзного государства.

Учитываю большую палитру возможностей, классификацию специализированной языковой модели Союзного государства можно предложить как некоторое комплексное решение, включающее такие позиции, как Mixture of Experts, мультимодальность, диалоговость, закрытость.

В качестве основы для разработки специализированной языковой модели Союзного государства можно предложить:

1. Формирование жанрово-стилистических доверенных русскоязычных и белорусскоязычных текстовых и устных корпусов.

Формирование жанрово-стилистические доверенных русскоязычных и белорусскоязычных текстовых и устных корпусов размером 50-100 ГБ (от 20 000 озвученных предложений различной длины и эмоциональной окраски), включающих: сбор текстов и аудиозаписей по направлениям: художественная литература отечественных писателей, культурно-исторические тексты, юридические и нормативные документы, медицинские публикации и др. Источниками данных могут являться национальные электронные библиотеки, официальные сайты государственных органов, тематические порталы.

Результат – предварительная фильтрация, нормализация и аннотирование текстовых данных.

2. Создание и поддержка расширенной базы данных на белорусском и русском языках для обучения специализированной языковой модели.

Разработка базы данных для обучения специализированной языковой модели имеет стратегическое значение для воспроизведения языковыми моделями социокультурных реалий, культурно-исторических контекстов и региональных различий народов Союзного государства, включая:

контроль оценочной окраски;

учет языковых и стилистических особенностей национальных языков, плохо представленных в общедоступных мультиязычных корпусах;

учет специализированной терминологии (законодательство, медицина, экономика, техника и др.);

соблюдение юридических норм, авторских прав по защите государственных и персональных данных;

реализацию требований о локализации данных и обеспечении прозрачности их использования.

Четкая методология сбора, фильтрации, аннотирования и хранения текстовых корпусов в сочетании с гибкой архитектурой хранилища и инструментами интеграции позволит получить высококачественный ресурс, готовый к использованию в современных NLP-конвейерах.

3. Глубокое обучение специализированной языковой модели на основе разработанной базы данных.

Ожидаемый результат – специализированная языковая модель, обеспечивающая качественную лингвистическую обработку информации на русском и белорусском языках, генерацию информации, релевантной российским и белорусским реалиям, и тематическую поддержку ключевых сфер общественной жизни.

Разработка специализированной языковой модели спланирована в рамках вновь создаваемой программы Союзного государства «Разработка цифровой платформы для решения приоритетных задач Союзного государства на основе технологий искусственного интеллекта, математического и суперкомпьютерного моделирования» в рамках отдельного задания. Предполагаемый заказчик от Российской Федерации – Министерство науки и высшего образования, ответственный исполнитель – научно-исследовательский вычислительный центр Московского государственного университета имени М. В. Ломоносова, от белорусской стороны – Национальная академия наук Беларуси, ответственный исполнитель – Объединенный институт проблем информатики НАН Беларуси.

Таким образом, можно сделать следующие выводы:

1. Разработка специализированной языковой модели, обученной на корпусе текстов, отражающих богатое наследие народов России и Беларуси (классическая литература, исторические документы, научные труды, фольклор), позволит создать ИИ-инструменты, которые будут транслировать и поддерживать общие для Союзного государства ценности.

2. Появление собственного ИИ-сервиса позволит гражданам Союзного государства, в первую очередь молодым гражданам, получать правдивую мировоззренческую информацию и решит важнейшие социальные и воспитательные задачи на общем пространстве. Среди них – сохранение культурного наследия, популяризация голосовых технологий и интерфейсов,

модернизация государственных сервисов, улучшение качества образования и здравоохранения, формирование мировоззрения подрастающего поколения в едином культурно-историческом пространстве народов России и Беларуси.

3. Специализированная языковая модель будет способствовать сохранению, развитию и популяризации языков народов России и Беларуси, станет мощнейшим инструментом для их продвижения и изучения на мировом пространстве.

4. Разработка специализированной языковой модели позволит обеспечить «цифровой суверенитет» в смысловом поле, где генерируемая информация будет соответствовать культурному коду народов Союзного государства.

Список использованных источников

1. Что такое трансформатор? // TRANSFORMER EXPLAINER. – URL: https://poloclub.github.io/transformer-explainer (дата обращения: 01.10.2025).

2. Выбор слоя активации в нейронных сетях: как правильно выбрать для вашей задачи. – URL: https://habr.com/ru/articles/727506 (дата обращения: 01.10.2025).

3. Reasoning-LLM: архитектура и обзор передовых моделей. – URL: https://habr.com/ru/companies/selectel/articles/892600 (дата обращения: 01.10.2025).

4. Котельников, А. Диффузионные модели / А. Котельников. – URL: https://education.yandex.ru/handbook/ml/article/diffuzionnye-modeli (дата обращения: 02.10.2025).

5. Мультимодальные языковые модели: как нейросети учатся видеть и слышать // Habr. – 2025. – URL: https://habr.com/ru/articles/892172 (дата обращения: 02.10.2025).

6. 10 лучших примеров генеративного ИИ. – URL: https://allsee.team/10-luchshih-primerov-generativnogo-ii (дата обращения: 02.10.2025).

7. Полное руководство – лучшие LLM с открытым исходным кодом для чат-ботов в 2025 году. – URL: https://www.siliconflow.com/articles/ru/best-open-source-LLMs-for-chatbots (дата обращения: 02.10.2025).

8. Полное руководство – лучшие LLM с открытым исходным кодом для кодирования в 2025 году. – URL: https://www.siliconflow.com/articles/ru/best-open-source-LLMs-for-coding (дата обращения: 02.10.2025).

9. Полное руководство – лучшие LLM для задач рассуждения в 2025 году. – URL: https://www.siliconflow.com/articles/ru/best-LLMs-for-reasoning-tasks (дата обращения: 02.10.2025).

10. Шесть больших языковых моделей с открытым исходным кодом. – URL: https://www.itweek.ru/foss/article/ detail.php?ID=228407 (дата обращения: 02.10.2025).

11. Использование внешних AI-моделей в Loginom. – URL: https://loginom.ru/blog/ai-models-loginom (дата обращения: 02.10.2025).

12. ЛИНЕЙКА: Каков реальный размер контекста ваших моделей длинного контекста языка? – URL: https://github.com/NVIDIA/RULER (дата обращения: 02.10.2025).

УДК 004.056

Подходы интеллектуальной ресурсной аналитики промышленно-коммунальных объектов

И. В. Войтов, В. Н. Штепа✉, А. В. Дубина
Белорусский государственный
технологический университет, Минск
E-mail: shtepa@bstu.by

Введение

Под ресурсной аналитикой организации можно подразумевать многофакторный процесс оценки и анализа внутренних ресурсов компании (материальных, трудовых, финансовых, технологических, информационных и ряда других) с целью определения их текущего состояния, эффективности использования и выявления скрытых резервов для достижения целей бизнеса, ситуационного прогнозирования. При правильной постановке задач и качественной реализации она помогает узнать, как объект использует свои потенциальные возможности для получения результата, и является основой для принятия стратегических решений, повышения конкурентоспособности и экологической нейтральности [1].

Можно выделить следующие детализированные подцели промышленной ресурсной аналитики в рамках общего тренда Индустрии 4.0 [2]: оценка текущего ресурсного потенциала, выявление резервов для повышения эффективности, определение сильных и слабых сторон компании, повышение конкурентности предприятия, улучшение экологической безопасности. При этом виды ресурсов, потенциально подлежащие анализу: финансовые (выручка, показатели ликвидности, платежеспособности, рентабельности); материальные (сырье, материалы, оборудование, здания); трудовые (персонал, его производительность и эффективность); технологические и информационные (технологические и информационные системы, инновационные сегменты); организационные (структура управления, бизнес-процессы).

Укрупненно можно выделить следующие этапы проведения такой аналитики на основе применения цифровых моделей [1, 2]: создание цифрового профиля (детальное описание и оценка всех имеющихся ресурсов исходя из применяемого (проектируемого) технологического регламента); оценка эффективности использования ресурсов; анализ финансового состояния (оценка ликвидности, устойчивости и рентабельности); комплексная оценка (определение общего показателя эффективности); экологиче-

ский анализ (выявление критических для экологической безопасности граничных значений параметров технологических процессов).

Вместе с тем информационно-аналитических систем, способных в оперативном режиме адаптироваться под изменения условий работы и обеспечивать заданное экономико-экологическое управление ресурсами (формировать профильные целевые показатели), фактически не внедрено. Соответственно, обоснование подходов и разработка программных решений мониторинга (контроль и прогноз) использования ресурсов является актуальной научно-практической задачей.

Последовательность ресурсной аналитики организаций (на примере использования технологической воды)

В рамках использования водных ресурсов на производстве можно выделить следующие подсистемы [3, 4]:

− транспортировка воды от источников (централизованная система водоснабжения или скважины) и ее водоподготовка;

− использование водных ресурсов в производственных процессах и ее загрязнение до состояния «сточная вода, требующая очистки»;

− сбор и аккумулирование воды в рамках ее водоотведения;

− очистка (при необходимости) сточной воды на локальных очистных сооружениях согласно нормативным требованиям установленных в рамках предельно-допустимых концентраций загрязнителей.

Необходимо акцентировать внимание на том, что на большинстве предприятий такая технологическая цепочка водопользования (под понятием подразумеваются любые варианты применения технологических водных растворов: мытье помещений, узлов и агрегатов; питье и отведение продуктов жизнедеятельности животных и птиц; охлаждение конструкций; растворение производственных реагентов и другие операционные действия) и водоотведения (включает редукцию поллютантов) опишет бóльшую часть производственных операций разноотраслевых объектов [5]. Таким образом, ресурсная аналитика применения технологической воды может рассматриваться как одна из ключевых при выполнении комплексных исследований в данном направлении (рис. 1 и 2).

Даже оценочный анализ сквозной цепочки водопользования (см. рис. 1) демонстрирует значительную сложность и фактическую невозможность построения единой динамической математической модели [5, 6]. Такой вывод основывается на том, что комплекс процессов водопользования и отведения включает параллельно-последовательное протекание биологических, физических и химических процессов, связанных с водными растворами – подавляющее большинство из них не имеет аналитического описания, не говоря уже об отсутствии аддитивности при таком моделировании.

Исходя из экспертного опыта [7], как один из вариантов проведения системного анализа предлагается итерационная последовательность (рис. 2). Она базируется на объединении при системном анализе, как теоретической основы создания информационно-аналитических продуктов мониторинга ресурсов промышленно-коммунальных объектов, технологических проектных решений и физико-математических моделей с целью синтеза обоснованного и адаптивного технического задания (которое не является стационарным, а корректируется в зависимости от внешних условий на протяжении всего жизненного цикла продукта) [8].

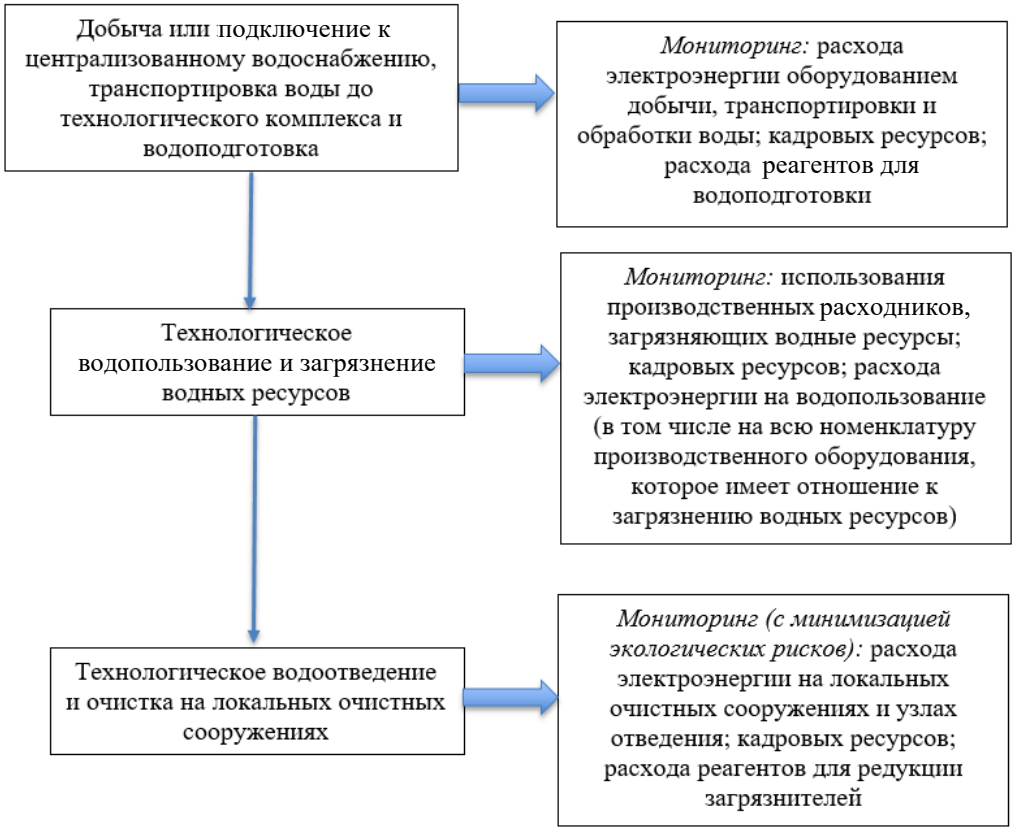


Рис. 1. Обобщенная схема сквозной ресурсной аналитики, связанной с водопользованием (от точки забора до точки отведения (сброса)

Практическая реализация интеллектуального модуля ресурсной аналитики

На основе методики (рис. 2) обосновано и создано программное обеспечение «Интеллектуальный модуль ресурсной аналитики» для деревоперерабатывающего предприятия, которое используется для поддержки принятия решений при производственном использовании технологической воды на основе интеллектуальных технологий (казуальные модели, теория

прецедентов, нечеткие нейронные сети, теория игр) мониторинга технологических процессов (рис. 3 и 4).

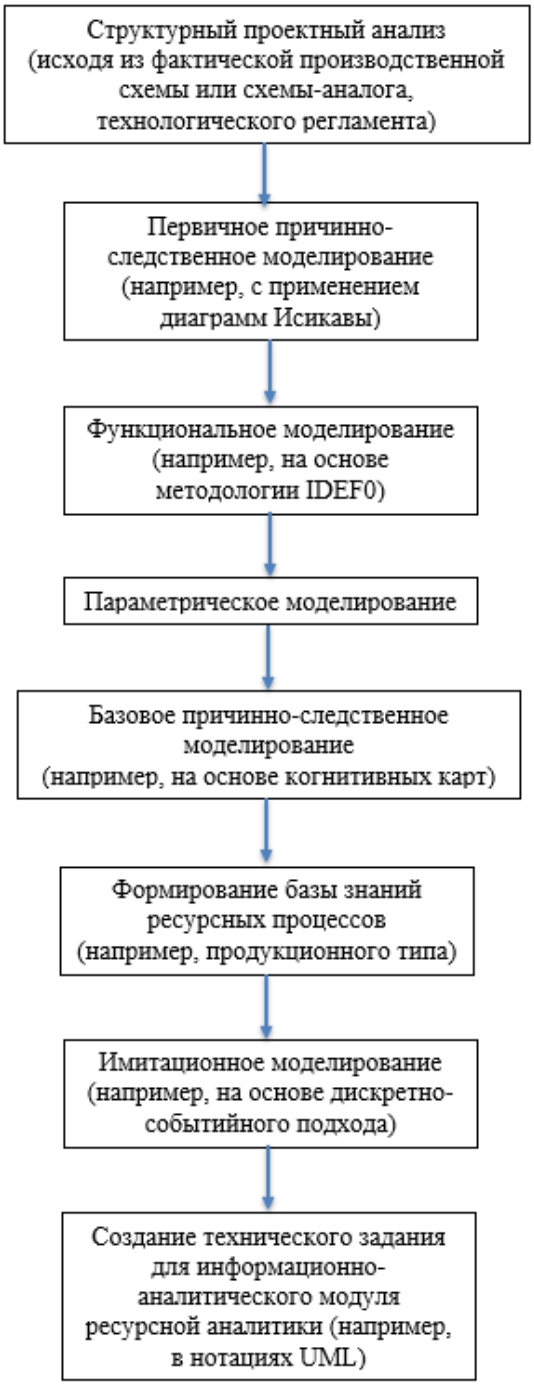


Рис. 2. Методика системного анализа комплексной ресурсной аналитики с конечной целью синтеза адекватного и адаптивного технического задания на продуктовые интеллектуальные решения (итерационные обратные связи не показаны, поскольку они при реальной работе могут иметь место между всеми этапами исследований)



Рис. 3. Схема информационных потоков интеллектуального модуля ресурсной аналитики

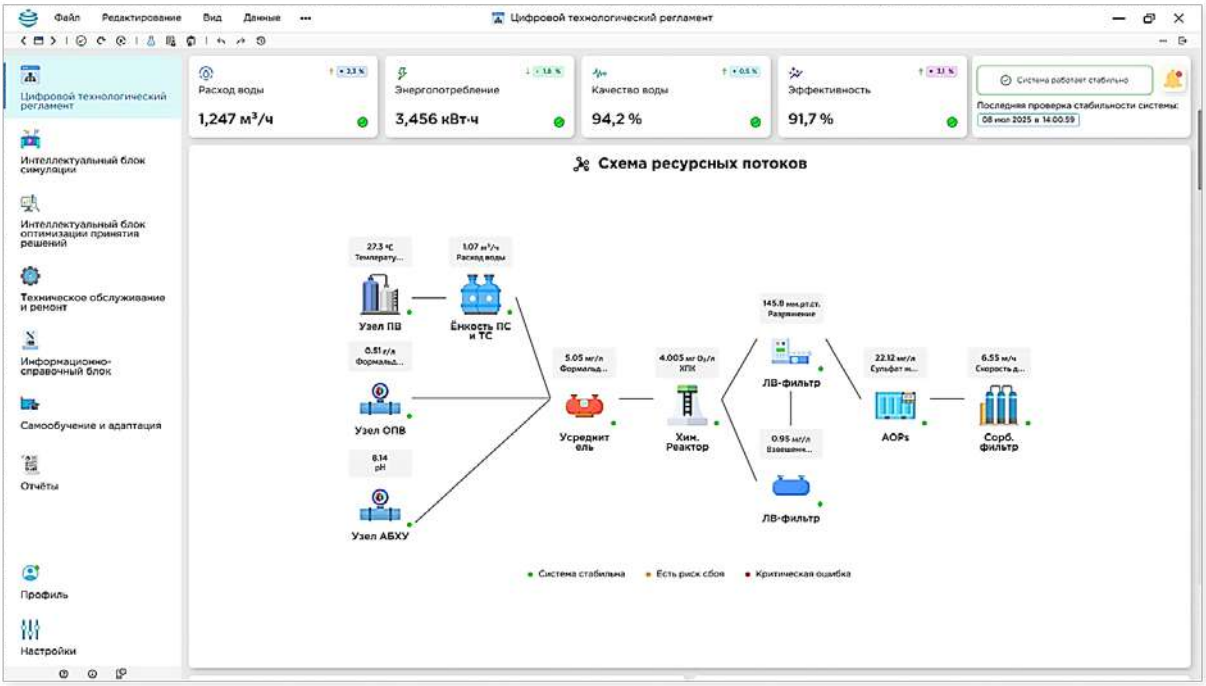


Рис. 4. Интерфейс ИМРА (на примере моделирования процессов водопользования, отведения и очистки формальдегидсодержащих сточных вод деревоперерабатывающего предприятия)

Основные технологии, на основе которых создана ИМРА: C# (язык программирования), .NET 8.0 LTS (основная платформа разработки), Avalonia UI 11.x (кроссплатформенный UI фреймворк (Windows, Linux, macOS), LiveCharts2 (библиотека для real-time графиков и визуализации данных); при этом для моделирования очистки сточных вод в наиболее

сложном узле (реакторе, использующем модифицированный процесс Фентона) применены нечеткие нейронные сети.

Заключение

В рамках имитационного моделирования функционирования ИМРА сформулированы основные потенциальные преимущества ее промышленного внедрения (в том числе в организациях водопроводно-канализационного хозяйства): повышение квалификации и переподготовка специалистов с использованием созданных цифровых решений без отрыва от производства; более точное и своевременное принятие управленческих решений, направленных на оптимизацию ресурсоэффективности; выявление и устранение проблем в бизнес-процессах; улучшение операционной ресурсной деятельности; повышение конкурентоспособности за счет более эффективного использования информации и ресурсов.

Дальнейшие исследования логично направить на формирование математических моделей, которые бы позволили реализовать самообучение используемых моделей.

Список использованных источников

1. Дубровский, А. В. Методические аспекты комплексного анализа и оценки эффективности использования ресурсного потенциала предприятия / А. В. Дубровский, Е. И. Щенникова // Экономика устойчивого развития. – 2023. – № 2. – С. 186–193.

2. Ильясова, М. К. Методические подходы к оценке ресурсного потенциала организации / М. К. Ильясова // Ученые записки Крымского инженерно-педагогического университета. – 2021. – № 4. – С. 103–109.

3. Штепа, В. Н. Интеллектуальная система анализа и прогноза экологической безопасности биологических сооружений очистки сточных вод / В. Н. Штепа, Н. Ю. Золотых // Первая выставка-форум IT-академграда «Искусственный интеллект в Беларуси» : сб. докл., Минск, 13–14 окт. 2022 г. / Объединенный институт проблем информатики Национальной академии наук Беларуси. – Минск : ОИПИ НАН Беларуси, 2022. – С. 41–45.

4. Alekseevsky, D. G. Formalization of the Task of Creating a Mathematical Model of Combined Wastewater Treatment Processes / D. G. Alekseevsky, Ye. Yu. Chernysh, V. N. Shtepa // Journal of Engineering Sciences: peer-reviewed scientific journal. – 2021. – Vol. 8, iss. 2. – P. H1–H7.

5. Штепа, В. Н. Цифровизация водопроводно-канализационного хозяйства с учетом требований экологической безопасности окружающей среды / В. Н. Штепа, Я. Ю. Ерш // Инжиниринг: теория и практика : материалы II Междунар. науч.-практ. конф., Пинск, 6 мая 2022 г. – Пинск : ПолесГУ, 2022. – С. 45–47.

6. Войтов, И. В. Использование цифровой платформы водоотведения для предупреждения технологических рисков биологической очистки сточных вод / И. В. Войтов, В. Н. Штепа // Инновационные биотехнологии для охраны окружающей среды: от теории к практике : материалы I Междунар. науч.-практ. конф., Минск, 23–25 апр. 2024 г. /

Институт микробиологии НАН Беларуси; организационный комитет конференции ; А. А. Шепшелев (пред.) и [др.]. – Минск : ИнМи, 2024. – С. 83–84.

7. Штепа, В. Н. Структура и функционал интеллектуальной системы поддержки принятий решений в водоотведении / В. Н. Штепа // Информатика и кибернетика : научный журнал. – 2022. – № 3(29). – С. 51–57.

8. Штепа, В. Н. Рациональные подходы к внедрению цифровых двойников в водопроводно-канализационные хозяйства / В. Н. Штепа // Информатика и кибернетика. – 2024. – № 3. – С. 51–57.

УДК 004.896

Классификация интеллектуальных инструментальных средств для автоматизации бизнес-процессов предприятия

С. Н. Генрихсен✉, **А. М. Седун**
Белорусский государственный
экономический университет, Минск
E-mail: diks80@mail.ru

Введение

Снижение цен на нефть и другие природные ресурсы изменило расклад сил в мировой экономике, которая подошла к новому этапу развития. Нельзя не согласиться с теми экспертами, которые указывают на необходимость перехода экономики к инновационной модели развития. Построение экономической модели, основанной на знаниях и инновациях, выбрано в качестве основного приоритета в реализации Национальной стратегии устойчивого развития Республики Беларусь [1, 2]. Одной из основных задач Национальной стратегии является построение интеллектуальной экономики. В документе отмечено, что в условиях глобальной борьбы за технологическое лидерство и исчерпания факторов экстенсивного роста экономики главным драйвером развития становится цифровизация. Цифровые платформы, роботизация и технологии искусственного интеллекта постепенно выходят за рамки производства, охватывая все сферы жизнедеятельности. Для достижения поставленных целей особое внимание уделяется внедрению информационных технологий нового поколения, роботизации, использованию облачных технологий, систем искусственного интеллекта в обработке большого объема данных [3].

В рамках деятельности экономических субъектов хозяйствования одним из перспективных направлений цифровизации бизнес-процессов организаций является использование такого инструментального средства, как RPA (RPA-Robotic Process Automation) [4, 5].

В отличие от классической автоматизации технология RPA подразумевает имитацию действий человека-пользователя (сотрудника) и даже их замену программными средствами.

Как правило, RPA используют графические и текстовые интерфейсы наподобие человека-пользователя. При этом они различают содержание установленных программных приложений, а также их отдельных элементов, таких как текст, таблицы, поля для ввода, кнопки и т. д.

RPA реализует часто повторяющиеся (рутинные), трудоемкие, хорошо алгоритмируемые операции, подпроцессы или бизнес-процессы.

Современные RPA начинают все шире использовать в своей деятельности также современные модели и методы искусственного интеллекта (ИИ) даже в пределах одной организации, все больше применяя модели проектирования и реализации интеллектуальных агентов.

Следует отметить, что в литературных источниках приводятся описания реализации различных RPA, а также их классификации по тем или иным критериям [3, 4]. Анализ показал, однако, что в настоящее время в литературе отсутствует единая принятая классификация PRA.

Как и в любой новой внедряемой технологии, классификация RPA может сыграть ключевую роль для руководства при выборе и дальнейшем эффективном внедрении и использовании RPA, а также направлений их развития.

Кроме того, классификация представляет в адекватном и удобном для пользователя виде рассматриваемые варианты реализации RPA и содержит в себе достаточно полную информацию о составляющих ее объектах.

Основная часть

В Докладе ЮНКТАД (Конференция ООН по торговле и развитию – United Nations Conference on Trade and Development, UNCTAD) за 2024 г. «О цифровой экономике» отмечается, что процесс стремительной цифровизации продолжается, трансформируя жизнь людей и их средства к существованию. В то же время нерегулируемая цифровизация рискует усугубить экологические и климатические проблемы. Подчеркивается настоятельная необходимость разработки устойчивых стратегий и инструментов на протяжении всего жизненного цикла цифровизации.

RPA как один из инструментов цифровизации позволяет быстро среагировать на изменения в бизнес-процессах организации или экономической системы в целом. Кроме того, RPA является достаточно универсальным программным инструментом, который может взаимодействовать практически с любыми информационными системами и приложениями.

В настоящее время для взаимодействия с другими информационными системами и человеком RPA используют различные программные интерфейсы, такие как:

UI based RPA (user interface) – пользовательский интерфейс. Он отвечает за дизайн, наполнение, композицию, оформление всех графических элементов.

API based RPA (application programming interface) – программный интерфейс приложения. Представляет собой набор правил и протоколов для создания и взаимодействия с другими программными приложениями.

Cognitive-based RPA – интерфейс искусственно интеллекта и машинного обучение для обработки неструктурированных данных. Cognitive based RPA осуществляет обработку естественного языка (Natural Language Processing, NLP), а также распознавание образов, позволяя RPA принимать решения на основе данных, адаптироваться к сложным сценариям и выполнять задачи, требующие суждений, подобных человеческим.

С учетом вышесказанного на рис. 1 схематично представлена авторская классификация RPA.



Рис. 1. Классификация инструментальных средств RPA

Рассмотрим классификацию инструментальных средств RPA по различным критериям.

1) В зависимости от области применения.

В системе *государственного управления* RPA могут выполнять следующие функции:

– прием, регистрацию, первичную обработку обращений граждан и юридических лиц, а также направление ответов заявителям;

– автоматическое размещение информации на сайтах государственных органов в части их касающейся;

– взаимодействие с государственной системой электронного документооборота;

– другое.

Область бизнеса обеспечивает устойчивое социально-экономическое развитие общества и напрямую зависит как от уровня активности предпринимательства, так и от уровня цифровизации.

В настоящее время RPA интенсивно внедряются на предприятиях различных форм собственности, относящихся к одной из групп субъектов хозяйствования: субъекты малого, среднего и крупного предпринимательства [5].

Предприятия малого и среднего бизнеса составляют преобладающее количество субъектов хозяйствования. Указанный сектор экономики является достаточно мобильным и легко адаптируется к часто изменяющимся рыночным условиям, поэтому развитие малого и среднего предпринимательства служит фундаментом для возникновения условий интенсивного экономического роста.

Так, предприятия малого и среднего бизнеса, в отличие от крупных предприятий, имеют более простые бизнес-процессы, используют для автоматизации и роботизации типовые программные платформы [5]. Все эти особенности обусловливают хорошие возможности для роботизации их бизнес-процессов. Для предприятий малого и среднего бизнеса не исключается полная роботизация их бизнес-процессов.

К примерам роботизации можно отнести следующие бизнес-процессы предприятий малого и среднего бизнеса:

– ведение бухгалтерского и финансового учета;

– ведение базы данных сотрудников;

– анализ рынка товаров и услуг (используя интернет-боты);

– ведение системы документооборота;

– другое.

В организациях крупного предпринимательства необходимо обратить внимание на объемность информации, сложность ее сбора, обработки и хранения, а также на большое многообразие отчетности, представляемой как в государственные органы, так и заинтересованным пользователям

этой отчетности. Поэтому в настоящее время полная роботизация бизнес-процессов таких организаций едва ли возможна.

Для организаций крупного предпринимательства RPA могут выполнять следующие функции:

– роботизацию отдельных частей бухгалтерского и управленческого учета;

– сбор и анализ информации из различных информационных источников (интернет-чат-боты), в том числе интеграция данных из этих источников;

– проверку благонадежности контрагентов;

– мониторинг тендерных площадок;

– другое.

В учреждениях образования RPA могут эффективно использоваться:

– в системе дистанционного обучения;

– в разработках учебных планов и программ;

– другое.

В области здравоохранения RPA могут выполнять следующие функции:

– обмен информацией с пациентом о записи на прием, процедурах, изменении либо отмене;

– отправку пациенту инструкции по подготовке к конкретным медицинским обследованиям, а также результаты обследований;

– ведение медицинских карточек;

– другое.

2) В зависимости от специфики взаимодействия с человеком.

Автономные RPA отличаются минимизацией (даже полным отсутствием) участия в их работе человека. Такие RPA работают самостоятельно на локальном сервере или в облаке. Выполнение необходимых задач RPA осуществляется по таймеру или при наступлении определенных событий-триггеров. При этом события-триггеры могут порождаться как бизнес-процессом, который обслуживает автономный RPA, так и внешними бизнес-процессами. Схема взаимодействия представлена на рис. 2.

Автономные RPA содержат команды, определяющие простые, предсказуемые и повторяющиеся бизнес-процессы, с которыми могут справиться самостоятельно. Одно из самых больших преимуществ автономных RPA – высокая рентабельность инвестиций за счет освобождения сотрудников и увеличения скорости обработки данных.

Примером автономных RPA может служить робот в бухгалтерии, обрабатывающий каждую ночь счета-фактуры и формирующий отчеты к началу рабочего дня, начисление заработной платы.

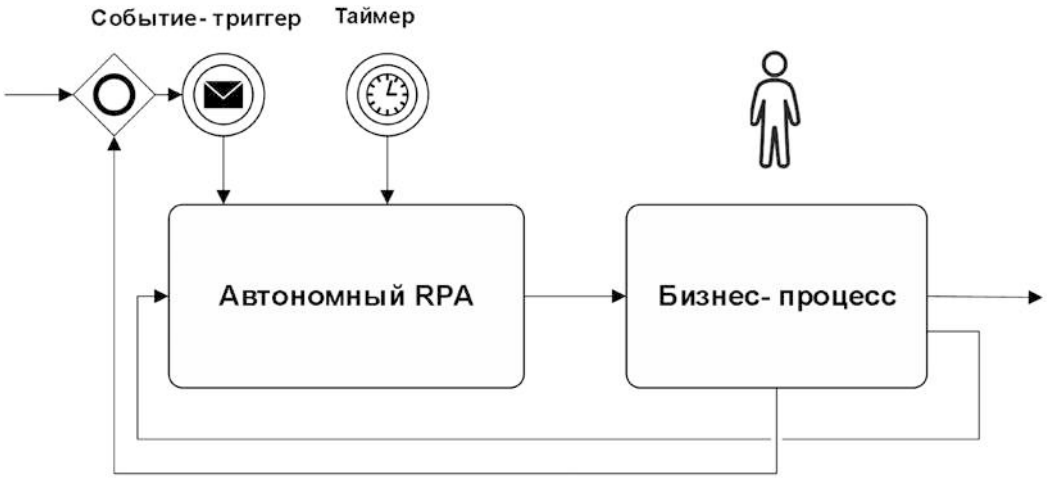


Рис. 2. Схема взаимодействия автономных RPA

Неавтономные RPA. В отличие от автономных взаимодействуют с человеком в режиме реального времени для выполнения сложных задач, требующих человеческого участия. Схема взаимодействия представлена на рис. 3.

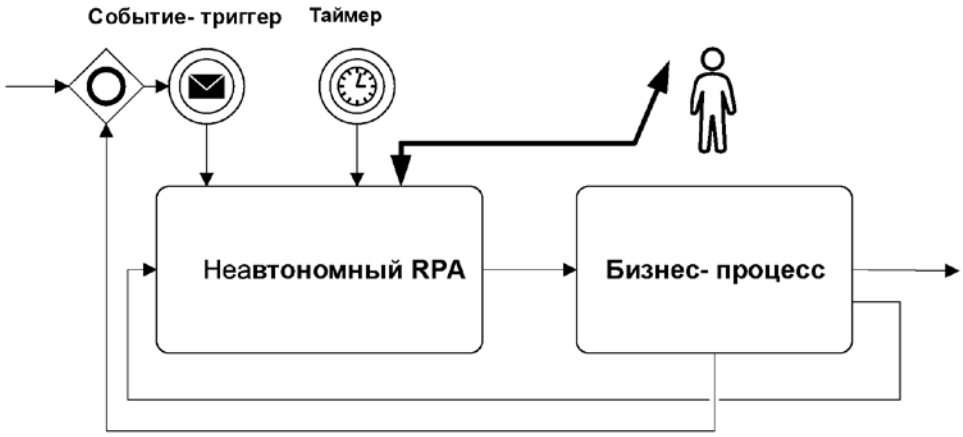


Рис. 3. Схема взаимодействия неавтономных RPA

Примером неавтономных RPA могут служить RPA, используемые для роботизации функционирования call-центра.

Гибридные RPA обладают определенной степенью гибкости, они сочетают в себе функции автономных и неавтономных RPA. Могут работать как автономно, так и во взаимодействии с человеком.

3) По специфике алгоритма функционирования.

Неадаптивные RPA работают на основе фиксированных, заранее разработанных алгоритмов без использования механизмов автоматического их изменения. При необходимости изменения алгоритма или используемых параметров RPA необходимо перепрограммировать. В настоящее время неадаптивные RPA наиболее широко представлены на ИТ-рынке.

Адаптивные RPA строятся на основе традиционных RPA, но дополняются механизмами (агентами) искусственного интеллекта (ИИ), используя машинное обучением (ML), а также аналитическими модулями обработки данных.

В литературных источниках [8–10] приводятся некоторые основные функциональные характеристики агентов ИИ, которые позволят их эффективно использовать для построения адаптивных RPA.

Использование свойств и алгоритмов работы ИИ-агентов при создании новых RPA позволит автоматизировать не только процессы, основанные на строгих правилах, но и более сложные, требующие осмысления данных, хранящихся в разных форматах, или их точного анализа.

Объединение технологий RPA и ИИ позволяет организациям роботизировать больше процессов, поскольку в этом случае рабочий процесс RPA более точно имитирует поведение человека, чем без использования ИИ-агентов.

Можно также отметить, что разработка *адаптивных RPA* является в настоящее время одним из перспективных направлений цифровизации во всех вышеприведенных областях деятельности.

Заключение

В условиях современного рынка одним из важных конкурентных преимуществ предприятия является способность быстрой адаптации к изменениям экономических условий ведения бизнеса и изменяющимся предпочтениям клиентов. Обеспечивает эту способность стремительное развитие цифровых технологий, автоматизация процессов операционной деятельности предприятия. Сегодня устойчивое социально-экономическое развитие напрямую зависит от уровня активности малого и среднего предпринимательства.

Организации малого и среднего предпринимательства в настоящее время составляют преобладающее количество субъектов хозяйствования. Данный сектор экономики является достаточно мобильным и легко адаптируется к часто изменяющимся рыночным условиям, поэтому развитие малого и среднего предпринимательства служит фундаментом для возникновения условий интенсивного экономического роста

Опираясь на анализ литературных источников, выделив наиболее часто используемые для классификации критерии, приведена авторская классификацию RPA. Классификация позволит систематизировать представление об RPA, понять их место в ИТ-инфраструктуре организации, целесообразность внедрения. Также следует отметить, что разработанная классификация может быть использована при разработке новых или модернизации существующих RPA.

Список использованных источников

1. Быков, А. А. Перспективы пост- и неиндустриального развития в условиях возможной трансформации системы международного разделения труда / А. А. Быков, А. М. Седун // Белорусский экономический журнал. – 2015. – № 2. – С. 4–23.

2. Быков, А. А. Иерархическая структура цепочек стоимости: статистический анализ по данным межрегиональных таблиц «затраты-выпуск» / А. А. Быков, А. М. Седун // Белорусский экономический журнал. – № 1. – 2021. – С. 45–58.

3. Национальная стратегия устойчивого развития Республики Беларусь до 2035 года / Мин-во экономики Респ. Беларусь. – URL: https://economy.gov.by/uploads/files/ObsugdaemNPA/NSUR-2035-1.pdf (дата обращения: 20.09.2025).

4. Сосновский, О. А. Программные роботы как инструмент цифровой трансформации экономики / О. А. Сосновский, С. Н. Генрихсен, А. М. Зеневич // Вестник БГЭУ. – 2023. – № 4. – С. 46–55.

5. Сосновский, О. А. Применение программных роботов для повышения эффективности бизнес-процессов предприятий малого и среднего бизнеса / О. А. Сосновский, С. Н. Генрихсен // Вестник БГЭУ. – 2024. – № 4. – С. 51–59.

6. Виды роботов RPA. – URL: https://unctad.org/system/files/official-document/der2024_overview_ru.pdf (дата обращения: 17.02.2025).

7. RPA инструменты и не только. – URL: https://habr.com/ru/articles/561270 (дата обращения: 17.02.2025).

8. Защищенный корпоративный маркетплейс и фабрика программных роботов. – URL: https://market.cnews.ru/articles/2023-07-28_zashchishchennyj_korporativnyj_marketplejs?erid=Kra23XMEy&p=homecnews (дата обращения: 17.02.2025).

9. Уатт, Дж. Машинное обучение: основы, алгоритмы и практика применения / Дж. Уатт, Р. Борхани, А. Катсаггелос ; пер. с англ. – СПб. : БХВ-Петербург, 2022. – 640 с.

10. ИИ-агенты и RPA. – URL: https://vc.ru/ai/1711436-ii-agenty-i-rpa-razbiraem-21-ih-razlichie-dlya-vybora-bolee-effektivnogo-podhoda (дата обращения: 17.02.2025).

УДК 004.82:16; 004.8.032.26; 547.9; 577.1

Применение генеративных нейронных сетей для дизайна новых ингибиторов белка KasA M. Tuberculosis

А. В. Гончар✉, **К. В. Фурс**, **А. В. Тузиков**
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: hanna.hanchar@gmail.com

А. М. Андрианов
Институт биоорганической химии
НАН Беларуси, Минск
E-mail: alexande.andriano@yandex.ru

Введение

Туберкулез по-прежнему остается одной из ведущих причин смертности в мире и представляет особую угрозу для пациентов с ВИЧ-инфекцией и сахарным диабетом [1]. Распространение лекарственной устойчивости существенно осложняет терапию, что делает особенно актуальными исследования по поиску новых противотуберкулезных средств. На ранних стадиях разработки все большее значение приобретают методы виртуального скрининга и машинного обучения, позволяющие ускорить поиск активных соединений и снизить затраты.

В числе перспективных терапевтических мишеней выделяется фермент KasA, играющий ключевую роль в биосинтезе жирных кислот клеточной стенки Mycobacterium tuberculosis. Потеря его активности приводит к лизису бактериальных клеток, что подтверждает критическое значение KasA для жизненного цикла возбудителя и определяет его как важную мишень для создания новых ингибиторов [2].

Цель настоящего исследования заключалась в создании генеративной модели нейронной сети для de novo дизайна малых молекул, потенциально активных против фермента KasA микобактерии туберкулеза. Для достижения этой цели были выполнены следующие этапы:

– разработка архитектуры нейронной сети для генерации низкомолекулярных химических соединений с высоким сродством к белку KasA;

– формирование обучающей библиотеки малых молекул, содержащих структурные элементы для селективного взаимодействия с активным центром фермента;

– обучение и тестирование модели на наборе соединений из созданной молекулярной библиотеки;

– генерация новых молекул, потенциально активных против белка KasA;

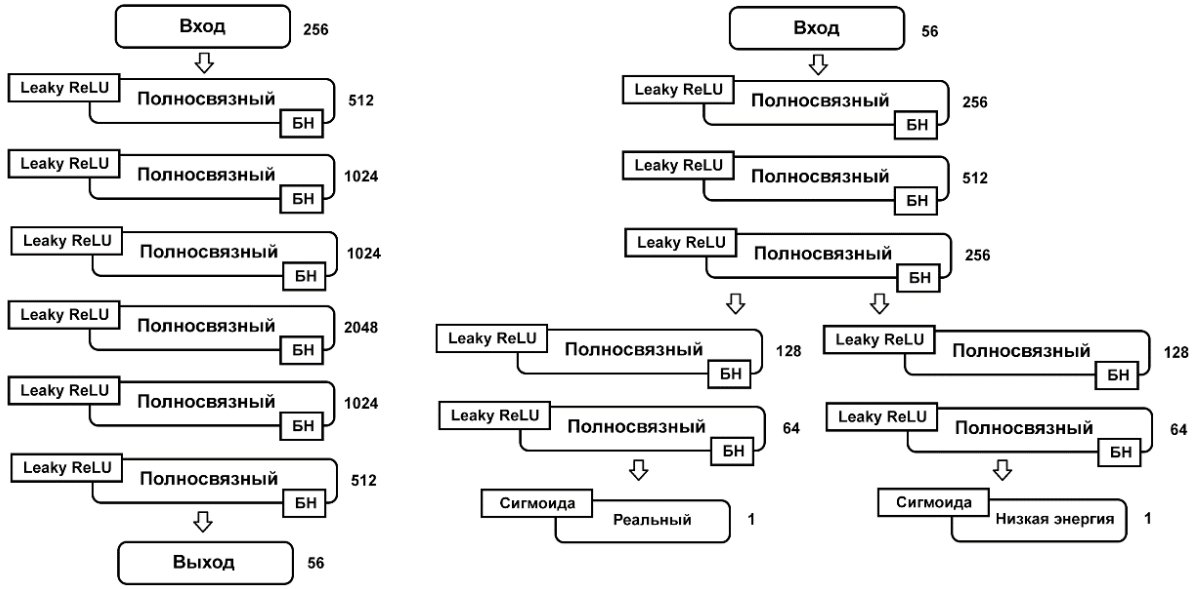
– молекулярный докинг сгенерированных молекул с целевым белком;

– молекулярная динамика перспективных соединений;

– оценка физико-химических свойств лучших соединений и отбор потенциальных ингибиторов МБТ.

Материалы и методы

Для решения поставленных задач была разработана модель генеративной состязательной нейронной сети с частичным привлечением учителя (SGAN, Semi-Supervised Generative Adversarial Network) (рис. 1), которая использует графовые эмбеддинги, извлеченные из латентного пространства вариационного автоэнкодера JTVAE (Junction Tree Variational Autoencoder) [3]. Обучение SGAN проводилось на наборе молекул из обучающей выборки с использованием в неявном виде значений энергии связывания с белком KasA. Молекулы в обучающем наборе данных разделяли на две группы, включавшие соответственно соединения с низкими (ниже – 8,2 ккал/моль) и высокими значениями энергии связывания с целевым белком, рассчитанными с помощью методов молекулярного докинга. Такая организация данных позволила модели генерировать новые соединения, структурно похожие на молекулы с высоким сродством к целевому белку.



а) *б*)

Рис. 1. Архитектура SGAN: *а*) генератор; *б*) дискриминатор

Формирование обучающего набора данных

Для формирования обучающего набора данных был проведен виртуальный скрининг библиотек Zinc15, ChemSpace и ChemDiv на веб-сервере Pharmit с использованием трех фармакофорных моделей, построенных на основе комплексов белка C171Q KasA с ингибиторами TLM и TLM5 (PDB ID: 4C6X и 4C72 соответственно). С помощью Python 3 и программного пакета RDKit (URL: https://www.rdkit.org) из отобранных молекул были удалены дубликаты и для каждого соединения получены канонические представления SMILES. В результате размер обучающей молекулярной библиотеки, сформированной с помощью фармакофорного анализа баз данных веб-сервера Pharmit, составил 58 815.

Для генерации графовых эмбеддингов, необходимых для обучения SGAN, из формата SMILES, полученного на этапе фармакофорного поиска, использовали предобученную модель нейронной сети JTVAE, которую разработчики обучали на выборке из 250 000 соединений, отобранных случайным образом в библиотеке Zinc15.

Для расчета значений свободной энергии связывания молекул из обучающего набора данных с целевым белком с помощью программы QuickVina2 был проведен молекулярный докинг этих лигандов с ферментом C171Q KasA в свободном состоянии. Структура белка извлекалась из комплекса C171Q KasA в кристалле с ингибитором TLM5 (ID PDB: 4C72). Структуры белка и лигандов были подготовлены к расчетам с помощью программного пакета MGLTools (https://ccsb.scripps.edu/mgltools/). Ячейка для докинга включала малонил-связывающий сайт KasA и имела следующие размеры: ΔX = 20,67 Å, ΔY = 24,8 Å, ΔZ = 16,46 Å с центром X = –7,24 Å, Y = –19,9 Å, Z = 6,75 Å. Значение параметра охвата конформационного пространства, определяющего широту поиска, было установлено равным 100.

Обучение SGAN

Для 58 815 малых молекул из сформированного обучающего набора данных с помощью сети JTVAE были получены графовые эмбеддинги, а затем эта виртуальная библиотека была разделена на две выборки – тренировочную и тестовую, состоящие из 47 052 и 11 763 векторов соответственно. Путем многочисленных экспериментов параметр, задающий количество эпох обучения, был выбран равным 150. Отношение частоты обучения генератора к дискриминатору было выбрано равным 0,3 к 0,7.

Функции потерь дискриминатора и генератора представлены в формулах

$$D_{loss} = BCE(D_{out1}(G(noise)), 0) + BCE(D_{out2}(real\_data), energy\_class), \quad (1)$$

$$G_{loss} = BCE(D_{out1}(G(noise)), 1), \quad (2)$$

где BCE – функция потерь «бинарная кросс-энтропия»; D – дискриминатор; D_{out1} – выход дискриминатора, отвечающий за предсказание реальности молекулы; D_{out2} – выход дискриминатора, отвечающий за предсказание класса энергии молекулы; G – генератор; noise – вектор гауссовского шума размерностью 256; real_data – вектор, соответствующий молекуле из тренировочной выборки; energy_class – 1 или 0 в зависимости от того, принадлежит молекула классу с низкой энергией связывания или нет.

Графики функций потерь генератора и дискриминатора SGAN представлены на рис. 2. Сплошная синяя линия соответствует потерям генератора, а красная пунктирная линия – потерям дискриминатора. Анализ графиков функций потерь дает основания полагать, что модель была обучена корректно, так как функции потерь сходятся, но не пересекаются.

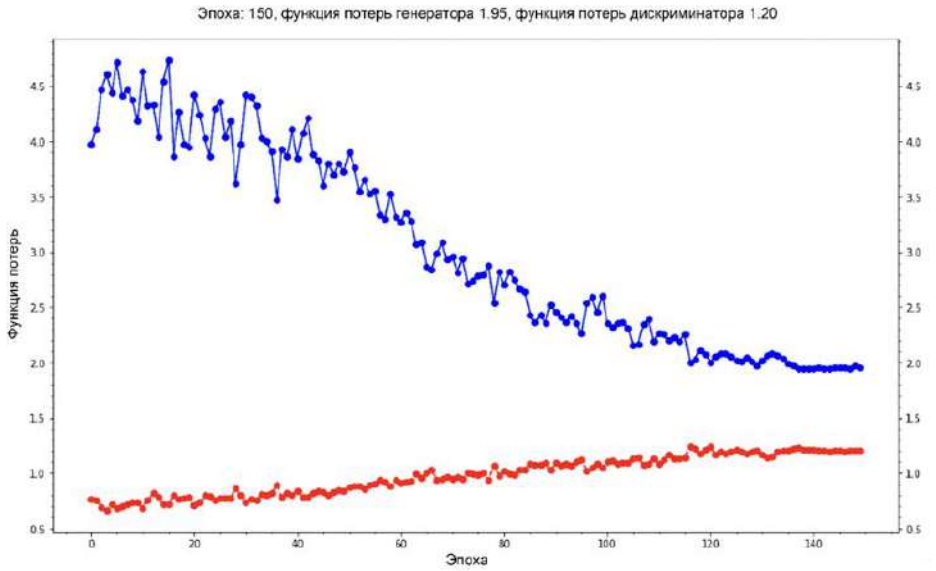


Рис. 2. Функции потерь генератора (синяя линия)
и дискриминатора (красная линия) SGAN

На основе случайных векторов длиной 256, имеющих гауссовское распределение, с помощью генератора SGAN было сгенерировано 200 000 векторов длиной 56 и получены соответствующие предсказания дискриминатора. Для оценки генеративных возможностей SGAN были выбраны два региона векторов. Первый (перспективный) регион включал векторы с вероятностью реальности выше 50 % и вероятностью низкого энергетического класса выше 50 %, а второй регион представлял область с вероятностью реальности молекул выше 50 % и вероятностью низкого энергетического класса ниже 50 %. Всего было получено 2565 и 2745 графовых эмбеддингов для первого и второго регионов соответственно, которые после

декодирования в JTVAE, удаления дубликатов и проверки корректности дали 1755 и 1882 уникальные молекулы.

Для выбора наиболее перспективных лигандов среди всех сгенерированных SGAN соединений (из обеих областей) значения свободной энергии связывания были переоценены с использованием оценочных функций RFScore-4 и NNScore 2.0. На основе предсказанных с помощью оценочных функций AutoDock Vina, RFScore-4, NNScore 2.0 значений энергии связывания для каждого соединения был рассчитан экспоненциальный консенсусный ранг (ECR) [4] и отобраны 20 лучших соединений для молекулярно-динамических расчетов.

Молекулярно-динамическое моделирование комплексов лиганд/белок проводили в воде с использованием Amber18 [5] и силовых полей Amber ff14SB (белок) и GAFF (лиганды). После минимизации, постепенного нагрева до 300 K, уравновешивания давления и достижения равновесия система моделировалась в ансамбле NPT при 300 K и 1 атм в течение 200 нс. В качестве контрольного соединения использовали ингибитор KasA TLM5 (K_d = 25,6 мкМ).

Свободную энергию связывания рассчитывали методом MM/GBSA для 40 комплексов, извлеченных из последних 80 нс траекторий (шаг 2 нс). Энтропийный вклад определяли с помощью Nmode из AmberTools19, а анализ траекторий проводили модулем CPPTRAJ.

Результаты

Для оценки потенциала разработанной модели нейронной сети были получены прогнозные показатели дискриминатора для 11 763 векторов из тестового набора данных, из которых в перспективную и вторую области вошли соответственно 3438 и 4448 молекулярных эмбеддингов. Для сгенерированных молекул и соединений из тестового набора был проведен молекулярный докинг с белком KasA с использованием того же вычислительного протокола, что и для соединений из обучающего набора данных. Результаты молекулярного докинга представлены в табл. 1.

Данные табл. 1 указывают на то, что дискриминатор SGAN может с высокой точностью предсказывать класс энергии на реальных (несгенерированных) данных, что подтверждает корректность обучения модели. Полученные результаты свидетельствуют о способности сети SGAN генерировать различные химические соединения с высоким сродством к белку KasA, особенно в перспективном регионе. Эти данные также указывают на то, что разделение молекул на классы на основе энергии связывания при обучении модели является эффективной стратегией для создания соединений с высоким сродством к целевому белку.

Таблица 1

Результаты молекулярного докинга для сгенерированных соединений и соединений из тестовой выборки

1 – перспективная область, 2 – вторая область		Количество соединений	Процент соединений с низкой энергией связывания (< –8,2 ккал/моль)	Средняя энергия связывания, (ккал/моль)
Сгенерированные соединения	1	1755	67	–8.5
	2	1882	40	–8.0
Соединения из тестовой выборки	1	3438	76	–8.7
	2	4448	21	–7.6

Анализ полученных траекторий молекулярной динамики позволил идентифицировать 12 соединений, у которых значения свободной энергии связывания (ΔG) оказались ниже, чем у контрольного соединения TLM5, рассчитанного по тому же вычислительному протоколу. Ниже представлены значения свободной энергии связывания и их стандартные отклонения по результатам молекулярной динамики для шести наиболее перспективных соединений (рис. 3), физико-химические параметры которых представлены в табл. 2. Учитывая стандартную ошибку метода MM/GBSA, составляющую около 2,9 ккал/моль, полученные данные указывают на то, что сродство этих соединений к KasA значительно выше по сравнению с ингибитором KasA TLM5. Также из табл. 2 видно, что отобранные соединения полностью соответствуют требованиям, предъявляемым к потенциальному лекарству согласно «правилу пяти» Липинского [6].

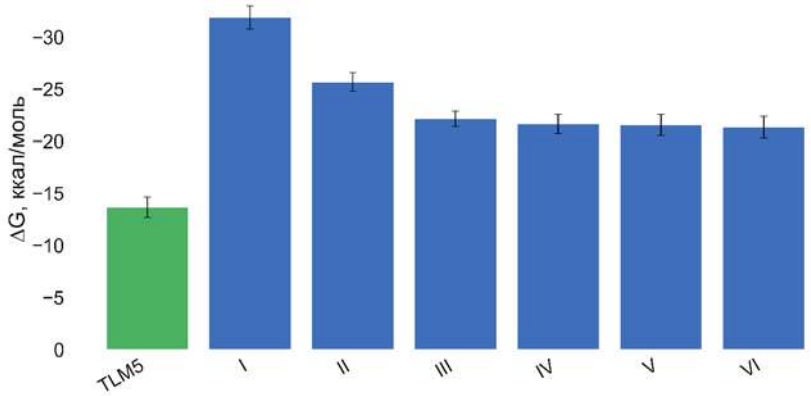


Рис. 3. Средние значения ΔG, рассчитанные для динамических моделей комплексов лиганд/KasA

Таблица 2

Физико-химические параметры идентифицированных соединений, связанные с «правилом пяти» Липинского [24], и их синтетическая доступность

Лиганд	Химическая формула	Молекулярная масса, Да	$LogP_{o/w}$	Синтетическая доступность	Число доноров водородной связи	Число акцепторов водородной связи
I	$C_{23}H_{29}N_3O_4S$	443.56	3.14	5.15	1	5
II	$C_{23}H_{30}N_6O_3$	438.52	2.58	4.22	0	6
IV	$C_{18}H_{22}N_6O_2S$	386.47	1.78	4.06	1	5
V	$C_{19}H_{14}FN_5O_2S$	395.41	2.81	3.47	1	6
III	$C_{23}H_{26}N_3O_3+$	392.47	2.69	3.8	2	5
VI	$C_{23}H_{35}N_5O_2$	413.56	3.22	4.95	1	5

Примечание: приведенные данные получены с помощью веб-сервера открытого доступа SwissADME (http://www.swissadme.ch); $LogP_{o/w}$ – липофильность соединения.

Заключение

В рамках исследования была разработана генеративная состязательная нейронная сеть (SGAN) с частичным привлечением учителя, адаптированная для поиска потенциальных ингибиторов фермента KasA, являющегося ключевым элементом биосинтеза клеточной стенки микобактерии туберкулеза.

Для реализации проекта была сформирована молекулярная библиотека соединений, содержащая низкомолекулярные аналоги известных ингибиторов KasA. В процессе исследования проведена оценка генеративных возможностей модели, выполнен молекулярный докинг сгенерированных соединений, а также оценка их сродства к мишени с использованием оценочных функций AutoDock Vina, RFScore-4 и NNScore 2.0. Проведенные молекулярно-динамические расчеты подтвердили стабильность взаимодействия идентифицированных соединений с ферментом KasA, что делает их перспективными кандидатами на дальнейшие экспериментальные исследования.

В результате были выделены шесть соединений, обладающих наилучшими характеристиками по сродству к ферменту и соответствующих требованиям фармакокинетики. Они могут стать основой для дальнейших экспериментальных исследований и разработки новых препаратов против лекарственно-устойчивого туберкулеза.

Работа выполнена при поддержке Консорциума и Программы по разработке портала по лекарственно устойчивому туберкулезу (URL:

https://tbportals.niaid.nih.gov) и поддержана Белорусским республиканским фондом фундаментальных исследований (проекты Ф23-007, Ф24КИТГ-016).

Список использованных источников

1. WHO. Global tuberculosis report 2024 / World Health Organization. – 2024. – URL: https://www.who.int/teams/global-programme-on-tuberculosis-and-lung-health/tb-reports/global-tuberculosis-report-2024 (date of access: 01.09.2025).

2. Identification of KasA as the cellular target of an anti-tubercular scaffold / K. A. Abrahams, Chun-Wa Chung, S. Ghidelli-Disse [et al.] // Nature Communications. – 2016. – Vol. 7. – Art. № 12581. – DOI: 10.1038/ncomms12581.

3. Jin, W. Junction tree variational autoencoder for molecular graph generation / W. Jin, R. Barzilay, T. Jaakkola // Proc. of the 35th Intern. Conf. on Machine Learning, Stockholm, 10–15 July 2018. – Stockholm, 2018. – P. 2323–2332. (Proceedings of Machine Learning Research).

4. Exponential consensus ranking improves the outcome in docking and receptor ensemble docking / K. Palacio-Rodriguez, I. Lans, C. N. Cavasotto, P. Cossio // Scientific Reports. – 2019. – Vol. 9. – Art. № 5142. – DOI: 10.1038/s41598-019-41594.

5. AMBER 2018 / D. A. Case, I. Y. Ben-Shalom, S. R. Brozell [et al.]. – San Francisco : University of California, 2018.

6. Experimental and computational approaches to estimate solubility and permeability in drug discovery and development settings / C. A. Lipinski, F. Lombardo, B. W. Dominy, P. J. Feeney // Advanced Drug Delivery Reviews. – 2001. – Vol. 46, № 1–3. – P. 3–26. – DOI: 10.1016/S0169-409X(00)00129-0.

УДК 004.822, 004.89

Тего-семантические основания искусственного моделирования интеллектуальной деятельности

А. Н. Гордей✉
Минский государственный
лингвистический университет, Беларусь
E-mail: alieks2001@yahoo.com

А. Н. Гордей, Р. С. Панащик, М. И. Святощик,
О. А. Стрельченок, В. В. Ткаченко
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск

А. Г. Шумилин
Национальная академия наук Беларуси,
ОФМИ, Минск

Введение

Синтаксический анализ (установление ролевой структуры предложения) и формализация семантики (определение ролевой структуры события, описываемого предложением) являются ключевыми этапами обработки и понимания текстов на естественном языке, и если синтаксический анализ (парсинг) активно применяется в компьютерных системах для формулирования правил или выявления противоречий в описании грамматики того или языка, то возможности программно-аппаратных комплексов в экспликации семантических отношений, формировании семантического представления текстов, кодировании и декодировании модели мира находятся в стадии становления. В то же время растущее количество научных статей и публикаций уже давно превысило тот порог [1], когда без автоматического анализа текстов, их аннотирования и реферирования, создания библиографических списков и электронных учебников «на лету» по запросу пользователя, автоматической семантической диагностики и экспертизы – другими словами, без искусственных интеллектуальных систем управления знаниями – не каждый ученый или специалист может успешно ориентироваться в своих и смежных научных областях. Активное развитие компьютерной техники требует новых подходов к обработке данных на естественном языке, новых теорий и алгоритмов, учитывающих смысловые значения языковых единиц, их контекстную и метаконтекстную зависимость. Ниже представлены принципы построения лексического анализатора

как инструментальной основы интеллектуальной системы семантической обработки текстов на базе теории автоматического порождения архитектуры знаний (ТАПАЗ-2) [2].

Искусственный интеллект в сфере обработки интернет-контента

Развитие искусственного интеллекта (ИИ) в области обработки данных на естественном языке идет по трем направлениям:

– интеллектуализация Интернета и веб-сервисов (проект Semantic Web 3.0 Консорциума W3C, проект международного общественного ресурса Scheme.org сообщества разработчиков Google, Microsoft, Yahoo и Yandex), превращение Интернета в «глобальный искусственный интеллект» (GGG) с одновременным представлением возможности подключения к нему автоматических систем управления (проекты «умный дом», «умный город» и т. д.);

– использование машинного обучения, нейронных сетей и иных алгоритмов статистической обработки контента, в том числе для оперирования большими данными (Big Data) при создании локальных и глобальных информационных ресурсов как основы цифровизации экономики и перевода ее на седьмой технологический уклад; самый известный пример такого подхода – разработка чат-ботов с генеративным ИИ (ChatGPT от OpenAI, DeepSeek и др.) как сервисов проверки и получения информации, призванных заменить «традиционные» поисковые системы;

– повышение эффективности работы больших языковых моделей (LLM) путем их интеграции с системой RAG (Retrieval-Augmented Generation) для подключения к релевантным внешним источникам знаний (статьи, технические описания и др.) и поиска информации.

Динамика развития деятельностно-ориентированной технологии (activity-based technology) с 2011 г., а также технологии Активного словаря (Activity Vocabulary) с 2017 г. показывает, что в ближайшие 10–15 лет главные усилия международных научных и финансовых центров будут сосредоточены на построении графов знаний для автоматического извлечения семантически значимой информации из поисковых страниц Интернета, т. е. на поэтапной разработке формализованного метаязыка, соотносимого по семантической мощности с естественным языком и способного в машиночитаемой форме представлять и преобразовывать знания [3].

Главным достижением Минской школы комбинаторной семантики является создание Теории автоматического порождения архитектуры знаний (ТАПАЗ-2, URL: http://tapaz.by/#/) для систем компьютерного моделирования интеллектуальной деятельности человека и решения изобретательских задач. Формальный аппарат ТАПАЗ-3, будучи финализацией Универсального семантического кода В. В. Мартынова, предназначен для

алгебраического кодирования семантики (значения знаков, смысла предложений, содержания стереотипов модели мира и связей в архитектуре знаний), т. е. исчисления семантики предметных областей, а также автоматической семантической разметки контента [4, с. 5–18].

Собственный код ТАПАЗ представляет собой неконвенциональное соединение алгебраического выражения с семантическим двойником (индивидом и его признаком в ролевой структуре события) и имеет строгие правила интерпретации типовых совмещений индивидов с переходом от морфологических категорий частей речи (ЧР) к семантическим категориям частей языка (ЧЯ) [5–7]. Таким образом, данный код кодирует все макропроцессы и процессы предметных областей, одновременно исчисляя их. Через обратный процесс декодирования стереотипов модели мира, а не категорий языковой картины мира код ТАПАЗ декодирует значения слов, словосочетаний, предложений и текстов, т. е. естественно-языковую семантику. В этом состоит его уникальность и преимущество перед упомянутыми мировыми разработками в данной области [8, с. 5–26].

Архитектура лексического анализатора

Идеологией ТАПАЗ является создание гибкой и открытой архитектуры многоуровневого универсального семантического анализатора контента для автоматической разметки корпусов текстов различных предметных областей (рис. 1). После разработки такой системы и ее обучения на размеченных в парадигме ЧЯ текстах из различных предметных областей появится возможность широкого внедрения и использования данной системы для интеллектуального управления знаниями.



Рис. 1. Архитектура универсального семантического анализатора

Ядром общей системы представляется лексический анализатор, с помощью которого по определенному алгоритму создается семантический словарь выделенной предметной области, в котором каждая лексическая единица (слово) тегирована определенным семантическим кодом-категорией из частной парадигмы ЧЯ [5, с. 173–179]. Тег – кодовая метка соответствующей категории в парадигме ЧЯ – *тайгенов (субстантивов) и ёгенов (предикативов)*, облегчающей процесс семантической разметки предложений и поиска необходимой информации. Например, слову *лектор* присваивается тег: *тайген, свернутый, сокращенный, физический, постоянный, нарицательный, качественный, одноместный*. Одному знаку также могут принадлежать разные теги: *столовая* (посуда), *столовая* (помещение для приема пищи) и т. д.; *он* – для обозначения как живых индивидов, так и неживых и т. д. Построенный таким образом тего-семантический словарь позволит решить проблему омонимии [9, с. 29–36] и будет использоваться в дальнейшем на этапе парсинга для синтаксического и семантического анализов текстов.

Алгоритм работы лексического анализатора

Алгоритм работы лексического анализатора состоит из следующей последовательности шагов (рис. 2):

1. Извлечение морфологических признаков ЧР (род, падеж, лицо, число, склонение и т. д.) каждого слова корпуса текста из общедоступных словарей.

2. Построение правил нечеткой логики анализа морфологических признаков ЧР и их классификации как ЧЯ.

3. Табличное однозначное определение семантических тегов слов (леммы и их словоформы) согласно правилам шага 2.

4. Ручная экспертная проверка табличного перевода и разметка определенных корпусов текстов (не классифицированных на шаге 2 и 3) для последующего обучения нейронной сети.

5. Обучение на базе правил и обучение нечетких множеств для распознавания частей языка с целью обобщения нечетких правил (шаг 2) на большие объемы текстов и решения проблем недостатка морфологических свойств и соответствующих правил.

6. Конечным этапом алгоритма является создание тего-семантического словаря (присвоение семантического тега каждой лексической единице предметной области).

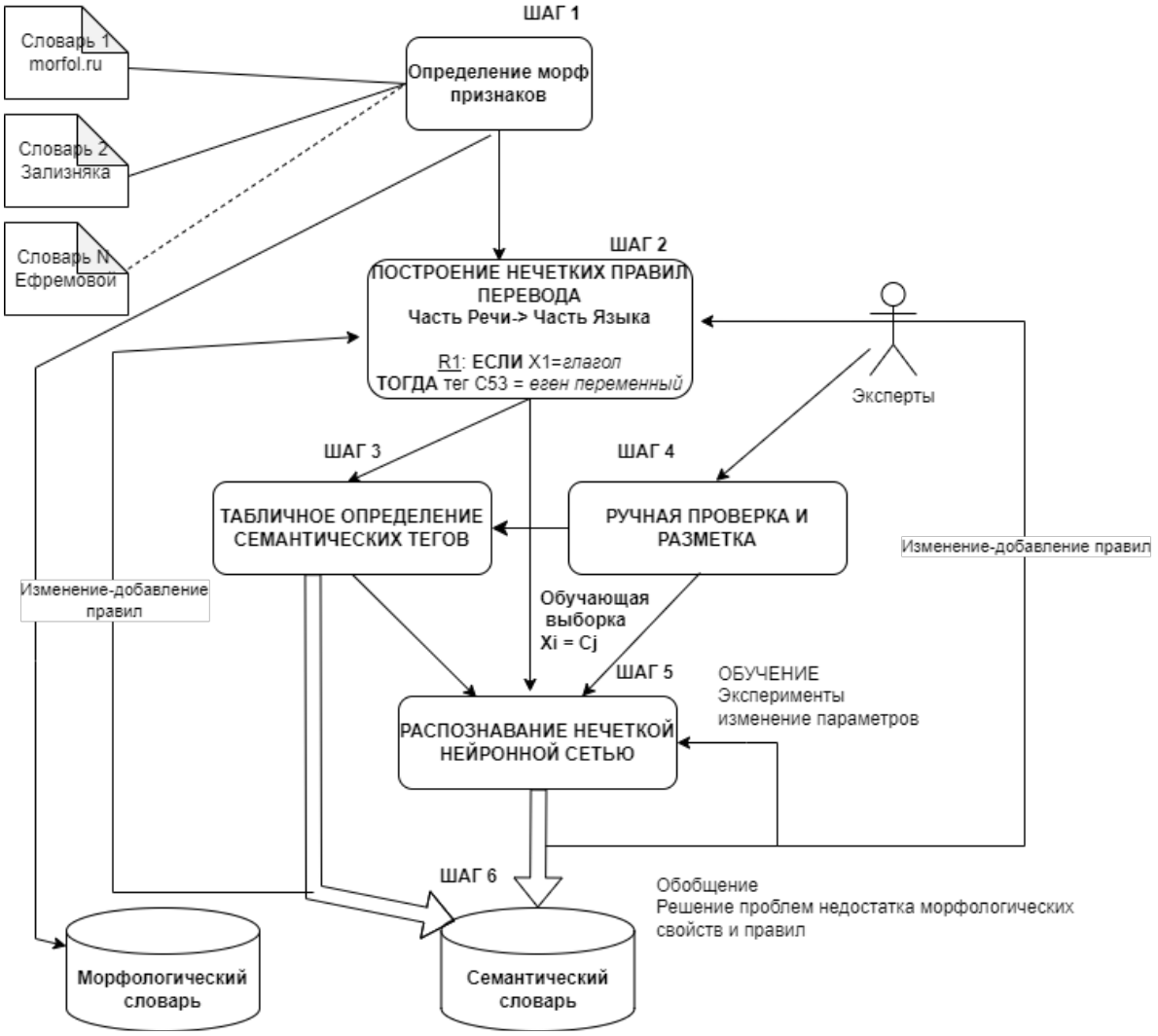


Рис. 2. Алгоритм семантической разметки и создания словарей

В качестве архитектуры выбрана нечеткая нейронная сеть NEFClass M (Modified Neuro Fuzzy Classifier), основанная на обобщенной архитектуре нечеткого персептрона с использованием алгоритма численной оптимизации гауссовской функции принадлежности и адаптированная под существующую задачу семантической разметки (рис. 3). Нечеткость возникает из-за неполных измерений свойств объектов (отсутствия или неполноты морфологических признаков слов в общедоступных словарях). Входным слоем сети является матрица ЧР и морфологических признаков, полученных из существующих словарей. Скрытым слоем является набор нечетких правил перевода ЧР и морфологических свойств в ЧЯ. Наконец, выходной слой сети представляет собой результирующий класс семантического тега по парадигме ЧЯ.

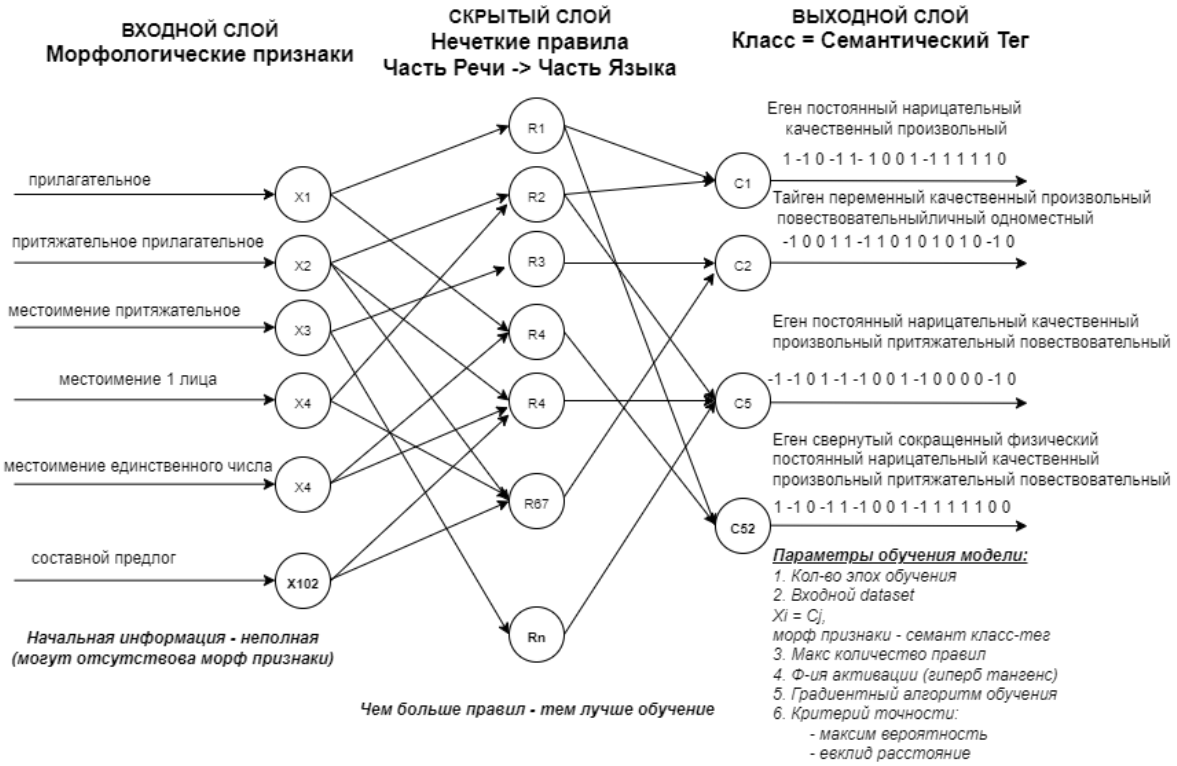


Рис. 3. Нечеткая нейронная сеть NEFCLASS

Нечеткие правила, описывающие данные, имеют следующую форму:

если x_1 является μ_1 **и** x_2 является μ_2 **и** … **и** x_n является μ_n,
то образец $(x_1, x_2, \dots, x_n)$ принадлежит классу i,

где $\mu_i, \dots, \mu_n$ – нечеткие множества.

Например,

если слово = *НИКТО*, как ЧАСТЬ РЕЧИ = *местоимение* **И** морф. признак = *отрицательное* **И** морф. признак = *единств. число*
то образец (*НИКТО*) имеет тег = *тайген сокращенный физический переменный качественный произвольный повествовательный неличный многоместный интенсивный* (код 1 0 0 0 0 -1 -1 0 1 -1 0 -1 0 0 0)

Реализация и результаты

Программа WEBSEMDICT, реализующая описанный метод, разработана как веб-приложение на основе микросервисной клиент-серверной архитектуры средствами ASP.Net Core MVC.

Дизайн и оформление сайта выполнены в соответствии со стандартами Bootstrap и некоторыми элементами CSS-3. На клиентской стороне также реализованы некоторые функции, написанные на языке JavaScript.

В качестве формы защиты персональных данных использован механизм аутентификации и авторизации Identity. В программе реализована функциональность разбиения пользователей по ролям, где к каждой роли прикреплена лишь часть функциональности приложения. Подтверждение пользователя идет по e-mail.

Общее количество слов, находящихся в базе, составляет около ~4 млн, что может говорить о достаточно большом объеме данных. Для сравнения «Словарь современного русского литературного языка» (17 томов) содержит 120 480 слов, электронные ресурсы Института лингвистических исследований РАН – около 5 млн слов. В базе данных находится более 20 таблиц, содержащих сведения о пользователях, сведения о файлах корпуса подключенных текстов выбранной предметной области, слов и словоформ из обработанных текстов с их морфологическими и семантическими характеристиками в отдельных таблицах. Механизм автоматической конвертации морфологических характеристик в семантические осуществляется с использованием 51 правила (без учета классификации по «жесткому тегу»). Для слов, словоформ, а также составных номинативных единиц, не подпадающих под названные правила автоматической разметки, в программе организован модуль ручной семантической разметки текста, а также модуль работы с нераспознанными словами из текстов, к которым относятся в том числе слова с опечатками. Семантические и морфологические характеристики любого слова поддаются изменению со стороны пользователей-экспертов, для которых определены соответствующие роли.

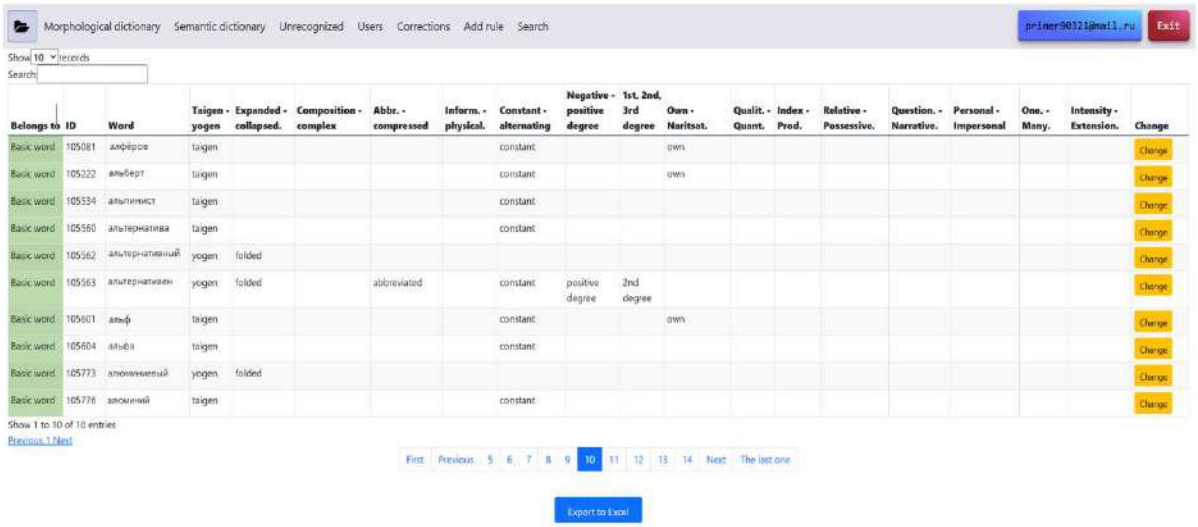


Рис. 4. Тего-семантический словарь

На данном этапе проекта реализована семантическая разметка уникальных слов из коллекции текстов предметной области «фотоника», состоящей из 81 текстового файла формата *.txt и *.docx с общим объемом

около 40 тыс. слов. В результате работы программы вместе с очисткой текстов от «шумов» создается таблица уникальных слов с их морфологическими характеристиками – морфологический словарь с разметкой слов по 32 признакам ЧР. Для этих же слов, которым присвоены уникальные индексы ID, программой создается таблица семантических характеристик – тего-семантический словарь с указанием их признаков как ЧЯ, всего 32 семантических признака или 360 уникальных тегов. Фрагмент тего-семантического словаря с элементами управления приложением со стороны клиента показан на рис. 4.

В дальнейшем планируется увеличение семантического словаря с загрузкой больших объемов текстов различных предметных областей и их автоматической обработкой с помощью нейронной сети совместно с разметкой на основе правил ЧР→ЧЯ.

Заключение

Новый подход к созданию интеллектуальных систем на базе ТАПАЗ для задач семантической обработки естественно-языковых текстов построен на классификации лексических единиц не по их частотно-статистическим характеристикам, как в алгоритмах генеративного искусственного интеллекта, а на основе их семантических отношений как ЧЯ и функций языкового выражения. Предложен и описан гибридный способ разметки предложений по ЧЯ с автоматизированной расстановкой тегов и построением тего-семантического словаря, в котором совмещено применение табличного однозначного тегирования ЧР-ЧЯ нечеткой нейронной сети и экспертной верификации.

Последующее масштабирование результатов проведенных исследований, а также задачи определения релевантности тренировочной области и подбора адаптивных алгоритмов машинного обучения опирается на всестороннее сотрудничество с ведущим белорусским специалистом в области искусственных нейронных сетей профессором В. А. Головко [11, с. 81–101], активное взаимодействие с профессором В. В. Голенковым в рамках платформы «Открытые семантические технологии проектирования интеллектуальных систем» (OSTIS Technology) [12, с. 81–98], которая ориентирована на разработку гибридных интеллектуальных систем. Уже есть хорошие примеры базы знаний OSTIS, описывающие лексические, синтаксические и семантические категории в терминах ТАПАЗ и доступные для реализации в алгоритмах обработки и понимания естественного языка [13, с. 183–188; 14, с. 123–140; 15, с. 192–221]. Технологии платформы OSTIS с ее оригинальным объектно-ориентированным языком программирования высокого уровня являются, на взгляд авторов, наиболее подходящими для погружения алгоритмов ТАПАЗ в программную среду.

Проведенная работа является существенным шагом в продвижении к созданию аппаратно-программного комплекса семантической обработки текстов на естественном языке, пониманию естественного языка и решению изобретательских задач в прямом и обратном направлениях от заданной конструктивно онтологии верхнего уровня к семантическому, синтаксическому и лексическому анализу любых технических и технологических описаний на естественном языке [16, с. 176–197]. Исследование вносит вклад в создание национальных искусственных интеллектуальных систем с опорой на комбинáторные методы или методы, основанные на знаниях.

Список использованных источников

1. Journal Rankings // SJR. Scimago Journal and Country Rank. – URL: https://www.scimagojr.com/journalrank.php?type=j (date of access: 21.12.2022).

2. Hardzei, A. Theory for Automatic Generation of Knowledge Architecture: TAPAZ–2 / A. M. Hardzei; transl. from Rus. I. M. Boyko. – Rev. English edn. – Minsk : Republican Institute of Higher Education, 2017. – 50 p.

3. Activity Vocabulary. W3C Recommendation. – URL: https://www.w3.org/TR/activitystreamsvocabulary/#dfn-activity (date of access: 18.04.2021).

4. Гордей, А. Н. Вновь о семантическом подходе к обработке данных на естественном языке / А. Н. Гордей // Иностранные языки в высшей школе / ФГБОУВО «Рязанский гос. ун-т имени С. А. Есенина». – Рязань : Редакц.-изд. центр РГУ им. С. А. Есенина, 2023. – Вып. 3(66). – С. 5–18.

5. Гордей, А. Н. Парадигма частей языка / А. Н. Гордей // Словообразование и номинативная деривация в славянских языках : материалы VIII Междунар. науч. конф., Гродно, 15–16 апр. 2003 г. / Гродн. гос. ун-т ; отв. ред.: С. А. Емельянова [и др.]. – Гродно, 2003. – С. 173–179.

6. Гордей, А. Н. Части языка и процедуры их разграничения / А. Н. Гордей // Пути Поднебесной. – Минск : БГУ, 2006. – Вып. 1. – Ч. 1. – С. 69–75.

7. Гордей, А. Н. Метасемантика языковых категорий / А. Н. Гордей // Вторые чтения, посвященные памяти профессора В. А. Карпова, Минск, 28 марта 2008 г. – Минск : Изд. центр БГУ, 2008. – С. 19–24.

8. Гордей, А. Н. Семантическая разметка события и ее отображение средствами китайского и русского языков / А. Н. Гордей // Иностранные языки в высшей школе / ФГБОУВО «Рязанский гос. ун-т имени С. А. Есенина». – Рязань : Редакц.-изд. центр РГУ им. С. А. Есенина, 2021. – Вып. 2(57). – С. 5–26.

9. Святощик, М. И. О конвертации частей речи в части языка / М. И. Святощик // Вестник МГЛУ. Филология. – Минск : МГЛУ, 2023. – Вып. № 1(122). – С. 29–36.

10. Алгоритм автоматизированной семантической разметки текстов / А. Н. Гордей, Р. С. Панащик, М. И. Святощик [и др.] // Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2024) : докл. XXIII Междунар. конф., Минск, 21 нояб. 2024 г. – Минск : ОИПИ НАН Беларуси, 2024. – С. 400–404.

11. Neuro-Symbolic Artificial Intelligence: Application for Control the Quality of Product Labeling / V. Golovko, A. Kroshchanka, M. Kovalev [et al.] // Communications in Computer and Information Science (CCIS). – Switzerland : Springer Nature Switzerland AG, 2020. – Vol. 1282. – P. 81–101.

12. From training intelligent systems to training their development tools / V. Golenkov, N. Gulyakina, N. Grakova [et al.] // Open Semantic Technologies for Intelligent Systems = Открытые семантические технологии проектирования интеллектуальных систем : сб. науч. тр. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: В. В. Голенков (гл. ред.) [и др.]. – Минск : БГУИР, 2018. – Вып. 2. – P. 81–98.

13. Hardzei, A. Perspective Approaches to Semantic Knowledge Representation and their Applications in the Context of the Task of Automatic Identification of the Se-mantically Equivalent Fragments of the Text Documents / A. Hardzei, Y. Krapivin // Open Semantic Technologies for Intelligent Systems = Открытые семантические технологии проектирования интеллектуальных систем : сб. науч. тр. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: В. В. Голенков (гл. ред.) [и др.]. – Минск : БГУИР, 2020. – Вып. 4. – С. 183–188.

14. Hardzei, A. Processing and understanding of the natural language by an intelligent system / A. Hardzei, M. Svyatoshchik, L. Bobyor, S. Nikiforov // Open Semantic Technologies for Intelligent Systems = Открытые семантические технологии проектирования интеллектуальных систем : сб. науч. тр. / Белорус. гос. ун-т информатики и радиоэлектроники ; редкол.: В. В. Голенков (гл. ред.) [и др.]. – Минск : БГУИР, 2021. – Вып. 5. – С. 123–140.

15. Hardzei, A. Universal Semantic Markup and Top-level Ontology Representation / A. Hardzei, M. Sviatoshchik, L. Bobyor, S. Nikiforov // Communications in Computer and Information Science (CCIS). – Switzerland : Springer Nature Switzerland AG, 2022. – Vol. 1625. – P. 192–221.

16. Hardzei, A. Plagiarism Problem Solving Based on Combinatory Semantics / A. Hardzei // Communications in Computer and Information Science (CCIS). – Switzerland : Springer Nature Switzerland AG, 2020. – Vol. 1282. – P. 176–197.

УДК 550.8.05

Управление экспериментальным нейроэлектрофизиологическим оборудованием с применением нейросетевых языковых моделей

А. А. Денисов
Белорусский государственный университет, Минск;
Институт физиологии НАН Беларуси, Минск
E-mail: an.denisov@gmail.com

Введение

В настоящее время в научной работе для обработки и анализа данных все шире применяются системы искусственного интеллекта (ИИ) на основе больших языковых моделей (Large Language Model, LLM) [1] – нейросетевых моделей, использующих глубокое обучение для обработки текста. Современные архитектуры на основе трансформеров и механизмов внимания позволяют LLM обрабатывать длинные текстовые последовательности и концентрироваться на важной информации во входных данных. Обученные на больших наборах, они успешно применяются для таких задач, как перевод, реферирование, ответы на вопросы, решение задач. Современные облачные сервисы на основе LLM также предлагают услуги по анализу и статистической обработке данных. Пользовательский интерфейс при этом часто реализуется в виде ИИ-ассистента, настроенного на помощи в решении определенных задач. Относительно новым направлением использования LLM в исследовательской работе является создание на их основе систем контроля и управления экспериментальным оборудованием. В работе [2] представлена разработка модульного ИИ-помощника для голосового взаимодействия человека со сложным научным оборудованием на основе масштабируемой системы с модулями LLM, специализирующихся на решении определенных задач. В работе [3] представлен подход по управлению сканирующим микроскопом с использованием сервиса ChatGPT. В таких случаях применяются LLM с десятками и сотнями миллиардов параметров. Для их работы требуются высокопроизводительные графические ускорители или облачные сервисы. Вместе с тем в научном оборудовании (особенно в портативном и автономном) часто применяются встраиваемые вычислительные системы с ограниченными ресурсами на основе микроконтроллеров и микрокомпьютеров, а использование удаленных сервисов

не всегда реализуемо. В данном случае возможно применение сокращенных (дистиллированных) версий нейросетевых моделей.

В лаборатории клеточной инженерии и нанобиотехнологий кафедры биофизики БГУ совместно с Центром мозга Института физиологии НАН Беларуси разработан комплекс аппаратуры и методик для регистрации электрической активности нейронов мозга крысы *in vitro*, реализующий управление оборудованием, обработку и анализ данных с применением микрокомпьютеров. Целью данной работы являлся анализ возможностей применения нейросетевых языковых моделей, установленных на микрокомпьютере класса Raspberry Pi 4, для управления процессом эксперимента.

Результаты и обсуждение

Общий вид автоматизированной установки для регистрации электрической активности нейронов срезов мозга крысы показан на рис. 1. Для регистрации полевых возбуждающих постсинаптических потенциалов и популяционных спайков используется 16-канальная система регистрации и стимуляции электрической активности с цифровым последовательным интерфейсом. Обработка, визуализация и запись регистрируемых сигналов, генерация последовательностей стимулирующих импульсов, управление аппаратурой, видеоотображение среза и микроэлектродов осуществляются при помощи микрокомпьютеров BeagleBone Black и Raspberry Pi 4 (4GB) под управлением программного обеспечения на языках C++ и Python.

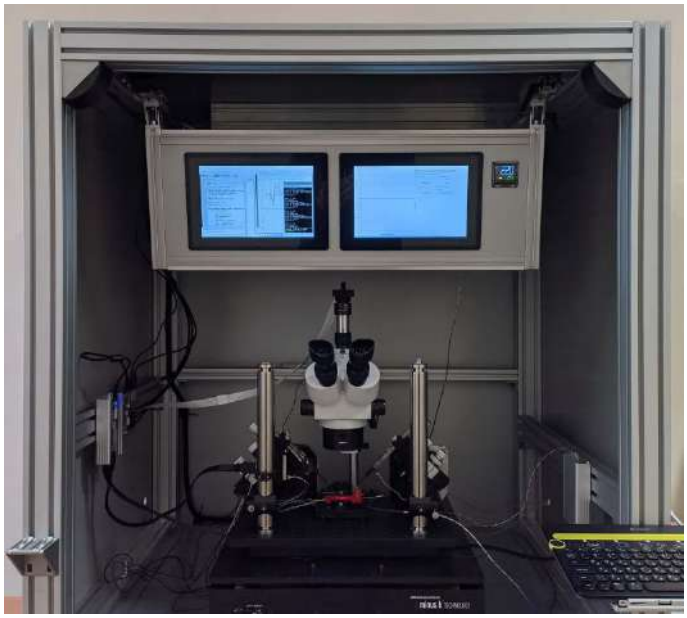


Рис. 1. Автоматизированная установка для регистрации электрической активности нейронов срезов мозга крысы

Система позволяет программно настраивать параметры встроенных фильтров для регулировки частотного диапазона регистрируемых сигналов, коэффициент усиления, выполнять аппаратное и программное устранение артефактов стимуляции. Применение микрокомпьютеров дает возможность гибко генерировать последовательности стимулирующих импульсов для исследования различных протоколов индуцирования синаптической пластичности, управлять температурой, временем и длительностью инфузии тестируемых препаратов в перфузионный поток и, таким образом, автоматизировать выполнение основных процедур в ходе эксперимента.

На рис. 2 показан фрагмент графического интерфейса для ввода параметров стимуляции и управлением регистрацией. Слева находятся классические элементы полей ввода для установки параметров амплитуды электрической стимуляции, продолжительности и частоты импульсов, запуска, приостановки и остановки регистрации. Справа реализовано окно ввода текстовой команды для изменения параметров с применением языковой модели.

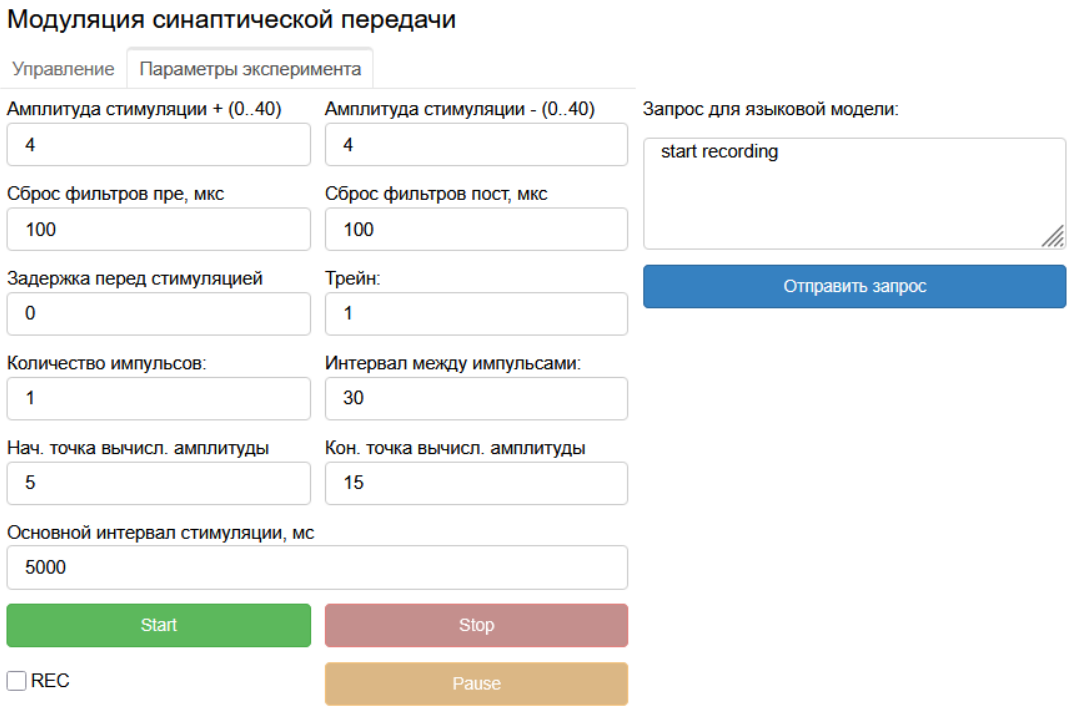


Рис. 2. Графический интерфейс для ввода параметров стимуляции и управлением регистрацией с применением классических полей ввода и с помощью команды для языковой модели

Полученная текстовая команда («user_input») далее встраивается в запрос, пример которого показан на рис. 3. В запросе формулируется роль ассистента, описываются необходимые действия и приводятся примеры

ответов. Затем на основе полученного от языковой модели ответа в формате JSON (если он корректно отформатирован) алгоритмически изменяются требуемые параметры.

```
"""
You are an intelligent assistant. Respond in JSON only.

If the request is about recording, then set "action": "recording", "state": start/stop/pause
If the requests is about stimulation amplitude, then set "action": "stimulation", "amplitude":
<amplitude value>.

Examples:
User: "start recording"
Response: {{"action":"recording","state":start}}

User: "stop recording"
Response: {{"action":"recording","state":stop}}

User: "set stimulation amplitude to 10"
Response: {{"action":"stimulation","amplitude":10}}


Process this user command:
User: "{user_input}"
"""
```

Рис. 3. Пример формирования запроса для языковой модели

Для обработки запросов были протестированы модели Deepseek coder 1.3b, Deepseek 1.5b, llama3.2 1b, Qwen2.5 coder 0.5b, Qwen2.5 1.5b, Qwen2.5 coder 1.5b, которые обладают размером, позволяющим запускать их на микрокомпьютере с объемом оперативной памяти, равным 4 ГБ. Наиболее высокую производительность (8,2 токенов/с) продемонстрировала модель наименьшего размера – Qwen2.5 coder 0.5b. Эта же модель показала и наихудшую точность, корректные ответы формировала только при командах, соответствующих примерам в запросе. Другие модели в среднем показывали производительность 3–4 токенов/с. llama3.2 1b отличалась относительно низкой точностью. Другие модели правильно формулировали ответ и в случае более произвольных формулировок, отличных от приведенных в примерах («baseline has stabilized, so begin recording», «set stim to 15»). Наибольшую производительность и точность показала модель Deepseek coder 1.3b, которая угадывала смысл команды, только косвенно относящейся к примерам («let’s go», «enough» – могла формировать корректные форматы запуска и остановки регистрации).

Заключение

Полученные результаты показывают, что модели типа «coder» достаточного размера, обученные в большей степени на примерах программного кода, справляются с задачей формализации запроса лучше, чем универсальные модели. Время обработки запроса для Deepseek coder 1.3b при

100%-й загрузке процессора составляло порядка 3–4 с, что достаточно для обработки команды в режиме типичного эксперимента по регистрации вызванных ответов, когда стимуляция и регистрация ответов производятся в среднем один раз в 20 с. Для ускорения обработки запросов возможна установка локальных встраиваемых ИИ-ускорителей.

Применение локальной нейросетевой языковой модели, установленной на микрокомпьютере, дает возможность использовать фразы естественного языка для управления процессом эксперимента. Одно из перспективных применений данного подхода – использование локальной модели для взаимодействия и обработки ответов других моделей, в том числе различных облачных сервисов, которые включают высокоуровневые варианты обработки данных. Например, при тестировании действия фармакологического препарата может производиться сравнение получаемых результатов с данными литературы по аналогичным экспериментам в режиме онлайн, предлагаться другие варианты проведения опытов. Интеграция такого «умного» экспериментального оборудования в глобальные ИИ-системы будет способствовать повышению производительности исследовательской работы.

Список использованных источников

1. Exploring the role of large language models in the scientific method: from hypothesis to discovery / Y. Zhang, S. A. Khan, A. Mahmud [et al.] // npj Artif. Intell. – 2025. – Vol. 1, no. 1. – P. 14.

2. VISION: a modular AI assistant for natural human-instrument interaction at scientific user facilities / S. Mathur, N. van der Vleuten, K. Yager, E. Tsai // Machine Learning: Science and Technology. – 2025. – Vol. 6, no. 2. – P. 025051.

3. Xie, Y. Toward Full Autonomous Laboratory Instrumentation Control with Large Language Models / Y. Xie, K. He, A. Castellanos-Gomez // Small Structures. – 2025. – URL: https://onlinelibrary.wiley.com/doi/10.1002/sstr.202500173 (date of access: 04.09.2025).

УДК 004.8

Искусственный интеллект в образовании: новые возможности для педагогической деятельности учителя

Ю. М. Дисковец
ГУО «Ордена Трудового Красного Знамени гимназия № 50 г. Минска»

Введение

В XX в. человечеству было представлено одно из наиболее значимых и полезных открытий – искусственный интеллект (ИИ). ИИ функционирует на основе принципов работы человеческого мозга, используя нейронные сети, в которых данные постоянно обрабатываются, изменяются, что приводит к обучению сети и формированию новых связей. В настоящее время нейронные сети широко применяются в различных сферах жизни и образование не является исключением. ИИ в образовании открывает новые горизонты для преподавателей. Инструменты ИИ для персонализации обучения позволяют:

– адаптировать учебный процесс к индивидуальным потребностям и возможностям каждого ученика;

– анализировать успеваемость, темп обучения и предпочтения;

– подбирать наиболее подходящие материалы, задания и стратегии обучения;

– использовать адаптивные обучающие платформы для автоматической корректировки уровня сложности заданий, обеспечивая оптимальную нагрузку для каждого ученика.

ИИ способствует автоматизации следующих рутинных задач:

– проверка грамматики, орфографии и пунктуации в письменных работах учащихся;

– освобождение времени учителя для более важных задач, таких как разработка креативных уроков, индивидуальная работа с учениками и повышение квалификации;

– помощь в организации учебного процесса, составлении расписаний и подготовке отчетов.

Улучшение вовлеченности учащихся может происходить через игровые обучающие платформы, использующие ИИ, чат-боты могут имитиро-

вать реальные диалоги, предоставляя учащимся возможность практиковать свои навыки говорения и аудирования в безопасной и комфортной обстановке.

Использование инструментов ИИ в образовательном процессе

Существует множество программ, которые могут помочь учителю в его деятельности, а также повысить интерес учащихся к процессу обучения. Современному учителю важно следить за научными достижениями, внедряя новейшие технологии в процесс обучения и воспитания.

В настоящей работе упор сделан на использовании инструментов, доступных без VPN, чтобы обеспечить максимальное удобство для всех.

Создание аудиозаписей

Развитие навыков аудирования – важная часть обучения английскому языку. ИИ может помочь в создании качественных аудиоматериалов.

Elevenlabs.io (URL: https://elevenlabs.io) представляет собой онлайн-сервис преобразования текста в речь, который не требует применения VPN (рис. 1). Этот ресурс используется для генерации аудиозаписей учебных текстов и диалогов. Процесс прост: вставляется текст, выбирается голос и скорость чтения, после чего получается готовый аудиофайл.

Например, создается аудиозапись текста о традициях и обычаях Республики Беларусь и учащимся дается задание прослушать информацию, после чего они должны ответить на вопросы. Этот метод позволяет активно обсуждать текст и разбирать сложные моменты.

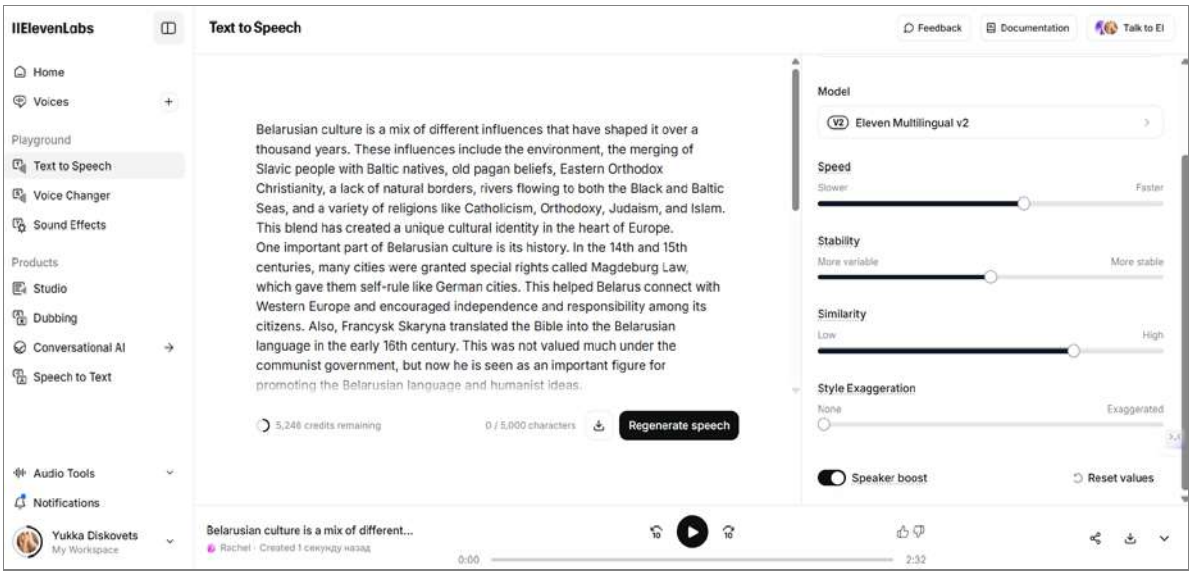


Рис. 1. Интерфейс сервиса elevenlabs.io

Генерация текстов

ИИ может помочь создавать тексты для различных целей: от примеров предложений до небольших рассказов.

Существует несколько аналогов ChatGPT. В своей деятельности автор использует *чат-бот Merlin* (URL: https://www.getmerlin.in/chat). *Чат-бот Merlin* – это усовершенствованный комплект чат-бота, который использует модели OpenAI и копирует интерфейс и функции ChatGPT (рис. 2). Для регистрации в чат-боте необходим адрес электронной почты. Чат-бот функционирует как расширение, которое устанавливается в браузер на компьютере. Бесплатная версия бота позволяет пользователю ввести 51 запрос в день. Также существует мобильная версия приложения на маркетплейсах.

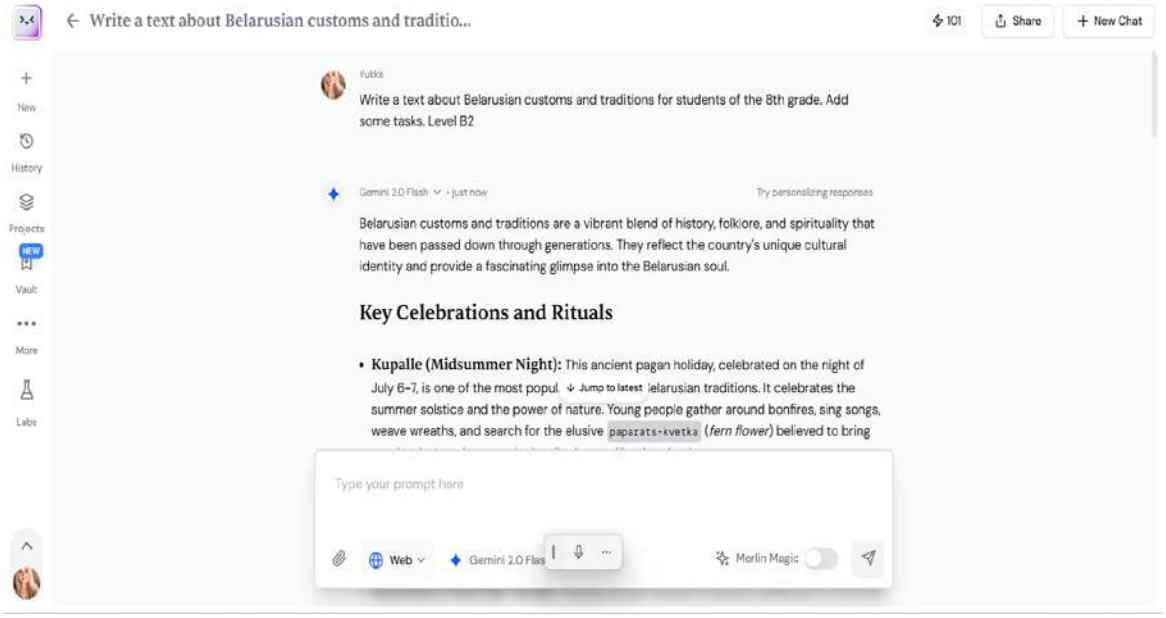


Рис. 2. Чат-бот Merlin

Чат-бот может генерировать несколько идей по определенной теме урока, создавать упражнения для отработки лексических и грамматических навыков, словарь-основу по теме урока, подборку идиом со значениями, примеры использования новых лексических и грамматических единиц в предложениях, перечень ресурсов для создания Flash-cards, тесты, открытые вопросы для обсуждения и другие типы заданий.

Автор использует возможности чата-бота Merlin для генерации текстов. Просто задает запрос, например: «Напиши текст о традициях и обычаях Республики Беларусь для учащихся 8 класса. Добавь задания. Уровень английского языка B1». Сгенерированный рассказ используется на уроке для чтения и обсуждения. Затем учащиеся пишут собственные рассказы, применяя новые слова и выражения.

Также сервис *Diffit* (URL: https://app.diffit.me/#topic) может сгенерировать статью на любую тему и вопросы по ней (рис. 3), придумать вопросы по интересующему видео или тексту.

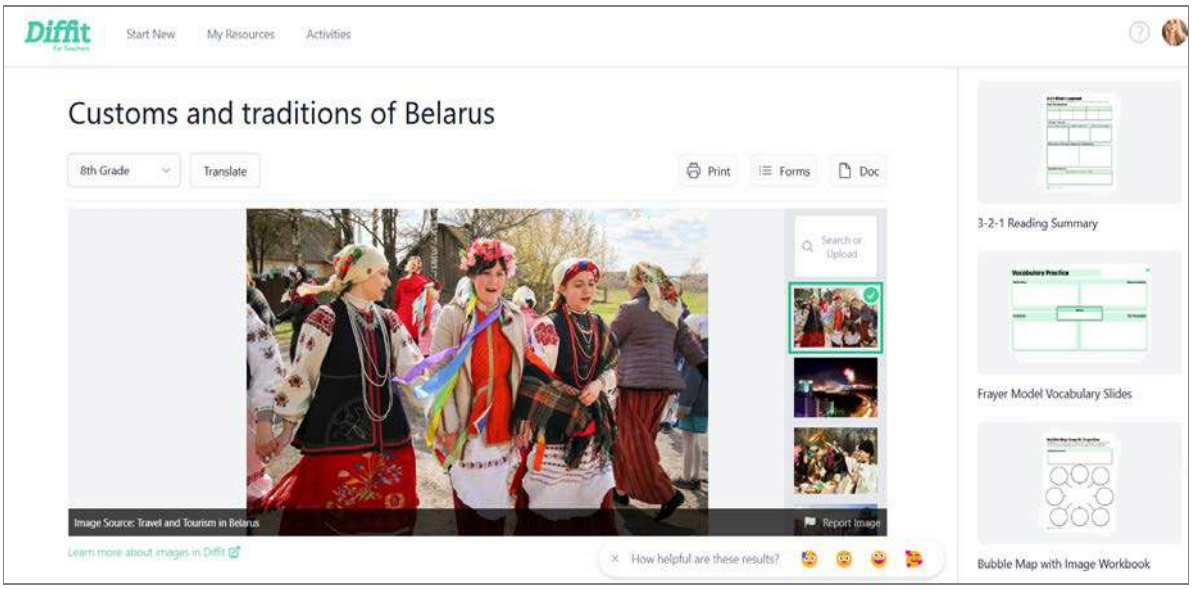


Рис. 3. Сервис Diffit

Создание заданий и тестов

ИИ также значительно упрощает процесс создания учебных заданий и тестов. Например, с помощью чат-бота Merlin можно запросить: «Сгенерируй 10 вопросов по теме Present Simple». Готовые вопросы используются для составления тестов в Google Forms, что облегчает оценку знаний учеников и выявление пробелов.

Создание картинок

Визуализация – критический аспект обучения. ИИ может помочь в создании изображений для уроков.

Bing Image Creator (URL: https://www.bing.com/images/create) – бесплатный инструмент на базе ИИ от Microsoft, который позволяет генерировать изображения из текстовых запросов. Используется для создания иллюстраций, которые затем включаются в презентации и задания.

Создание презентаций

Инструмент *Slidesgo* (URL: https://slidesgo.com) – это платформа, предоставляющая шаблоны, совместимые с PowerPoint и Google Slides, которая позволяет быстро создавать информативные и визуально привлекательные презентации. Просто вводится текст, и сервис автоматически генерирует слайды с заголовками, списками и изображениями. Это значительно экономит время на подготовку наглядных материалов для уроков.

Написание планов-конспектов

Чат-бот Merlin также обеспечивает возможность эффективного планирования уроков. Можно задать запрос, например: «Составь план-конспект урока по теме «Hobbies» для 6 класса, рассчитанный на 45 минут», и получить структуру, которая адаптируется под нужды учащихся.

Заключение

Современный мир технологий предлагает учителю большое количество возможностей разнообразить учебные занятия, а также облегчить подготовку к ним. Приведенный список инструментов ИИ – это всего лишь малая часть из того, что существует в цифровом мире. Использовать их или нет зависит только от желания и возможностей учителя. Как известно, именно развитие и движение вперед приводят к прогрессу, в том числе прогрессу в обучении и воспитании подрастающего поколения.

УДК 598.2:591.582.2:004.934

Автоматическая система распознавания голосовых сигналов птиц, основанная на технологиях искусственного интеллекта

Д. А. Жалова✉, О. В. Дыдо, В. А. Хохлов,
А. А. Бакунович, Е. А. Журавлева
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: daryazhalova@gmail.com

Введение

В последние годы наблюдается значительная конвергенция технологий искусственного интеллекта (ИИ) и биоакустики, что расширяет возможности мониторинга и охраны дикой природы. Одним из наиболее перспективных направлений является разработка автоматизированных интеллектуальных систем распознавания голосовых сигналов птиц для точного обнаружения и классификации их вокализаций. Птицы играют важнейшую роль в поддержании экологического баланса, выполняя функции опылителей, распространителей семян и индикаторов состояния окружающей среды. Поэтому возможность мониторинга их популяций с минимальным вмешательством человека становится все более актуальной [1].

Традиционные методы мониторинга птиц, такие как ручное наблюдение и экспертный анализ, требуют значительных временных затрат и являются трудоемкими. Эти подходы также подвержены человеческим ошибкам, что может негативно сказаться на точности собранных данных. В условиях, где необходимо быстро реагировать на изменения в экосистемах, такие методы становятся недостаточно эффективными. В то же время современные методы, основанные на алгоритмах ИИ, такие как глубокое обучение со сверточными сетями и архитектурами трансформеров, значительно повышают скорость и точность обработки данных [2]. Эти технологии позволяют системам распознавать сложные закономерности в криках птиц, адаптироваться к различным акустическим условиям и различать виды по голосам. Например, алгоритмы глубокого обучения могут анализировать большие объемы акустических данных, выявляя уникальные характеристики вокализаций, которые могут быть неочевидны для человеческого восприятия. Использование возможностей ИИ в мониторинге птиц не только упрощает процесс идентификации и отслеживания, но и открывает

новые горизонты для масштабируемых решений в реальном времени. Это особенно важно в сложных и динамично меняющихся природных средах обитания, где традиционные методы могут оказаться неэффективными. Немаловажно и то, что автоматизированные системы работают в условиях, где человеческое вмешательство затруднено, например в удаленных или труднодоступных районах [3].

Таким образом, в статье представлена усовершенствованная модель распознавания голосовых сигналов птиц белорусской фауны в реальных условиях [4]. В качестве базовой архитектуры для модели классификации используется предварительно обученная сверточная нейронная сеть (CNN) EfficientNetB3, модифицированная путем удаления конечного слоя классификации. Модель включает в себя несколько этапов предварительной обработки аудиосигнала и использует специализированную схему классификации, разработанную для задач с несколькими метками, что позволяет независимо определять 116 видов птиц в аудиосегменте.

Описание модели

EfficientNetB3 – это одна из моделей семейства EfficientNet – группы CNN. Изначально разработанная для задач классификации изображений EfficientNetB3 обеспечивает высокую точность при меньшем количестве параметров и меньших вычислительных затратах. Применительно к распознаванию голосов птиц типичный подход заключается в преобразовании аудиосигналов (птичьих криков и песен) в частотно-временные представления (например, спектрограммы или Mel-спектрограммы) и последующей обработке их как «изображений», передаваемых в CNN [5].

EfficientNetB3 была разработана с использованием метода поиска нейронных архитектур (Neural Architecture Search, NAS), что позволяет находить оптимальные конфигурации для различных задач. В ней используется метод составного масштабирования, который позволяет одновременно увеличивать глубину, ширину и разрешение сети. Это помогает достичь лучшего баланса между данными параметрами, что, в свою очередь, улучшает производительность модели. EfficientNetB3 включает в себя мобильные инвертированные узкие места (Mobile Inverted Bottlenecks), которые помогают уменьшить количество параметров и вычислительных затрат, сохраняя при этом высокую производительность.

Выбор такой архитектуры основан на ее высокой эффективности с точки зрения точности и сложности вычислений. Сочетая увеличение глубины, ширины и разрешения входного изображения, EfficientNetB3 обеспечивает конкурентоспособную производительность при умеренном количестве параметров. Более того, модели, построенные на базе EfficientNet, неизменно демонстрируют первоклассные результаты в различных зада-

чах, связанных с акустической классификацией птиц, включая соревнования по акустическому распознаванию видов птиц. Этот опыт подтверждает его практическую пригодность для подобных применений, что делает его предпочтительным выбором для данного исследования.

Записи вокализаций птиц предварительно обрабатываются путем преобразования необработанных сигналов в спектрограммы (или Mel-спектрограммы) – двумерное представление со временем на одной оси и частотой на другой, по сути, создавая «изображение», которое отражает распределение энергии по времени и частоте. Затем размер изображений спектрограмм изменяется до соответствующего входного разрешения, ожидаемого EfficientNetB3. Она обрабатывает изображения с помощью слоев сверток, где на ранних слоях обычно извлекаются низкоуровневые элементы (края, текстуры), а на более поздних слоях фиксируются высокоуровневые паттерны (частотные модуляции, гармоники), которые могут соответствовать определенным паттернам вокализации птиц.

EfficientNetB3 обладает рядом преимуществ по сравнению с другими архитектурами CNN в процессе распознавания голоса, а именно:

Эффективность: благодаря сбалансированной архитектуре EfficientNetB3 обладает высокой вычислительной эффективностью и при этом обеспечивает конкурентоспособную точность. Это особенно полезно при развертывании систем распознавания голосов птиц на периферийных устройствах или при проведении полевых исследований, где вычислительные ресурсы могут быть ограничены.

Устойчивость к изменениям: вокализация птиц часто содержит незначительные изменения во времени и частоте. Надежность и способность EfficientNetB3 иерархически фиксировать локальные и глобальные закономерности делают ее подходящей для различения разных видов или даже обнаружения вариаций в вызовах одного и того же вида.

Масштабируемость: в зависимости от размера и разнообразия набора данных EfficientNetB3 можно масштабировать (или даже комбинировать с методами расширения данных), чтобы повысить точность распознавания сложных видов птиц или шумного фона.

Компоненты EfficientNetB3

Основными компонентами архитектуры EfficientNetB3 являются мобильный инвертированный блок *MBConv* и оптимизация *Squeeze-and-Excitation*.

Инвертированный остаточный блок начинается с расширения входной карты признаков в более многомерное пространство с использованием свертки 1×1, затем применяет свертку по глубине в этом расширенном пространстве и, наконец, использует еще одну свертку 1×1, которая проецирует

карту признаков обратно в пространство с меньшей размерностью, такое же, как и входное измерение (рис. 1). «Инвертированный» аспект исходит из этого расширения размерности в начале блока и сокращения в конце, что противоположно традиционному подходу, где расширение происходит к концу остаточного блока.

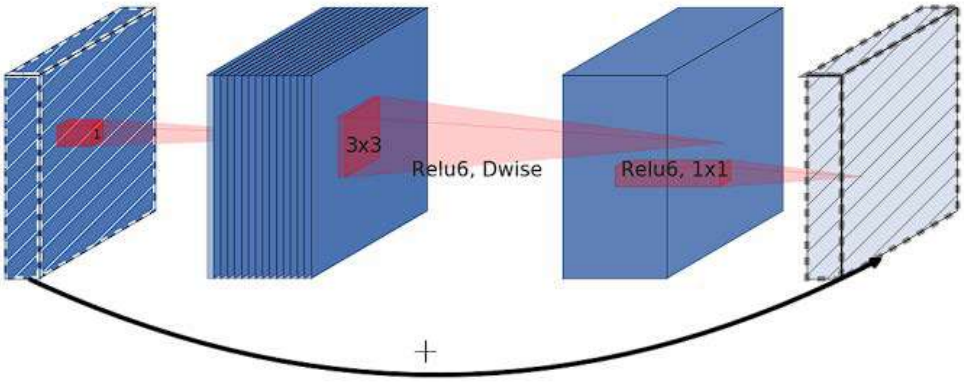


Рис. 1. Инвертированный остаточный блок

Squeeze-and-Excitation (SE) позволяет модели подчеркнуть полезные признаки и подавить менее полезные. Это делается в два этапа:

1. Squeeze (сжатие). Фаза объединяет пространственные измерения (ширину и высоту) карт признаков по каждому каналу в одно значение, используя глобальное усреднение. Это приводит к компактному дескриптору признаков, который суммирует глобальное распределение для каждого канала, сводя каждый канал к одному скалярному значению.

2. Excitation (возбуждение). На этом этапе модель, использующая полносвязный слой, применяемый после этапа сжатия, создает коллекцию весов по каналам (активаций или оценок). Последний этап заключается в применении этих изученных оценок важности к исходной карте входных признаков по каналам, эффективно масштабируя каждый канал по его соответствующей оценке.

SE – это строительный блок для CNN, позволяющий улучшить взаимозависимости между каналами путем выполнения динамической перекалибровки по всем каналам. Это означает, что, вместо того чтобы придавать всем каналам одинаковый вес, сеть будет динамически назначать высокий вес наиболее важным каналам. На рис. 2 показаны компоненты данного устройства. Этот процесс позволяет сети подчеркивать более важные признаки и преуменьшать значение менее важных, динамически корректируя карты признаков на основе изученного содержания входных изображений.

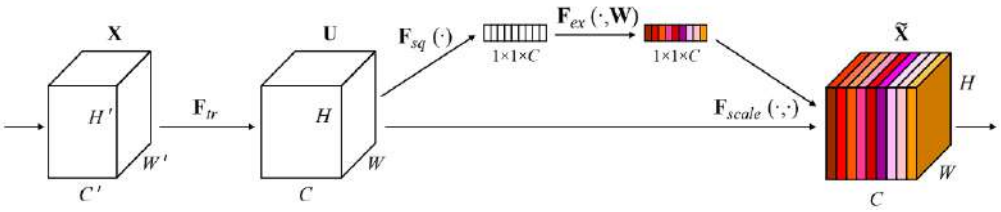


Рис. 2. Squeeze-and-Excitation-блок

EfficientNet включает в свою архитектуру все вышеперечисленные элементы, представленные на рис. 3. Размер входного изображения составляет 3x3 (RGB-изображение). Модель делится на несколько стадий, каждая из которых включает в себя различные слои свертки:

Stage 1 – слой свертки (Conv) с ядром 3x3 – отвечает за начальную обработку входных данных;

Stage 2 – слой MBConv (Mobile Inverted Residual Block) с ядром 3x3 – использует глубинные свертки и остаточные соединения для повышения эффективности;

Stage 3 – слой MBConv с ядром 3x3 – продолжает извлечение признаков из входных данных;

Stage 4 – слой MBConv с ядром 5x5 – позволяет захватывать более сложные паттерны за счет увеличения ядра;

Stage 5 – слой MBConv с ядром 5x5 – осуществляет дальнейшее извлечение признаков;

Stage 6 – слой MBConv с ядром 5x5 – углубляет модель для повышения точности;

Stage 7 – слой MBConv с ядром 5x5 – завершает процесс извлечения признаков перед выходом.

На выходе каждой стадии формируется карта признаков, которая передается на следующую стадию. Это позволяет модели эффективно обрабатывать информацию и извлекать значимые характеристики.

На выходе сверточной части сети добавлен слой глобального усредненного пулинга (GlobalAveragePooling2D), который выполняет агрегирование пространственной информации по каждому признаковому каналу. Применение данного слоя позволяет уменьшить размерность признакового пространства и способствует снижению риска переобучения, одновременно сохраняя наиболее релевантные признаки.

Следующим компонентом модели является слой *dropout* с вероятностью исключения нейронов, равной 0.4. Этот метод регуляризации используется с целью повышения обобщающей способности модели: случайное «отключение» части нейронов в процессе обучения препятствует избыточному запоминанию тренировочных данных и способствует формированию более устойчивых признаков.

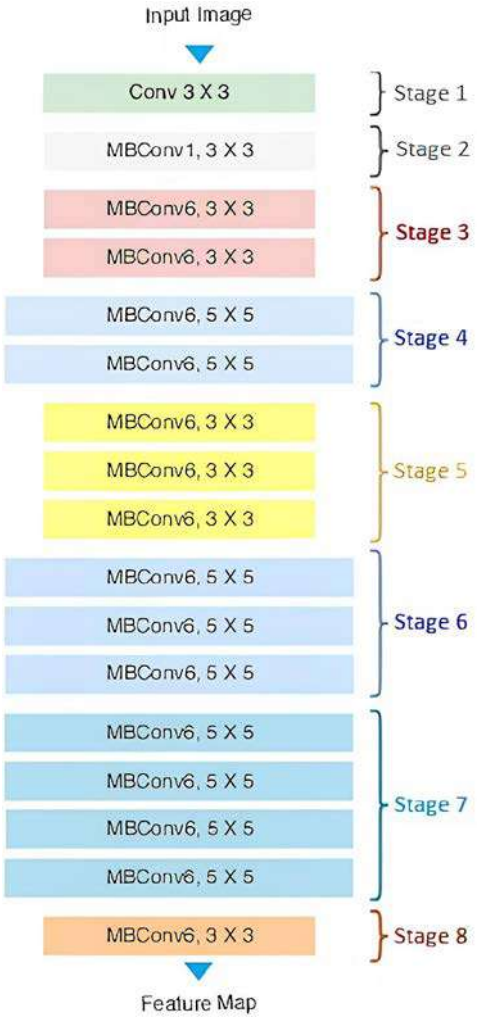


Рис. 3. Архитектура EfficientNet

Финальный полносвязный слой содержит 116 выходных нейронов с функцией активации *sigmoid*. Количество выходов соответствует числу классифицируемых классов – видов птиц. Выбор сигмоидной активации объясняется постановкой задачи в форме multilabel-классификации, при которой каждый класс (вид) предсказывается независимо от других. Это позволяет учитывать ситуации, когда в одном и том же аудиофрагменте могут одновременно присутствовать вокализации нескольких видов.

В качестве функции потерь используется бинарная кросс-энтропия (BinaryCrossentropy), что также соответствует специфике multilabel-задачи: каждый выходной нейрон интерпретируется как независимое бинарное решение (присутствует (не присутствует) вид птицы). Для оптимизации параметров модели применяется алгоритм *Adam*, сочетающий преимущества адаптивной коррекции градиента и импульсного обновления, что обеспечивает быструю сходимость и устойчивость при обучении на сложных и шумных данных.

Предварительная обработка данных

В ходе сбора и аннотирования голосовых сигналов животных в Беларуси специалистами-орнитологами была собрана обширная база данных (датасет) из различных источников. Она содержит более 2500 аудиозаписей 116 видов птиц, что дает приблизительно 42 000 аннотированных вокализаций [6].

Перед отправкой аудиозаписей на вход модели аудиозаписи проходят предварительную обработку – серию шагов, направленных на повышение качества извлекаемых признаков и улучшение результатов классификации. Первоначально удаляется фоновый шум, чтобы свести к минимуму влияние несущественных акустических помех, которые могут исказить спектральные характеристики целевого сигнала и снизить точность распознавания. Этот шаг улучшает соотношение сигнал / шум, что особенно важно для полевых записей, которые часто содержат различные фоновые звуки, такие как шум ветра, воды или жужжание насекомых. Затем записи делятся на фиксированные двухсекундные сегменты. Такой размер окна обеспечивает баланс между наличием достаточной информации и предотвращением перегрузки модели слишком большим количеством данных. Кроме того, чтобы расширить набор обучающих данных и повысить устойчивость модели к временным изменениям звука, используется стратегия расширения: границы каждого окна случайным образом сдвигаются на 0,5 с. Такой подход обеспечивает более разнообразное представление входных данных и расширяет возможности обобщения модели.

После сегментации каждое звуковое окно подвергается нормализации и динамическому сжатию. Нормализация устраняет различия в амплитуде между различными звуковыми фрагментами, гарантируя, что модель не будет сильнее реагировать на более громкие звуки. Сжатие усиливает сигналы низкой интенсивности, сохраняя тонкие, но важные акустические характеристики.

Наконец, из обработанных окон генерируются Mel-спектрограммы с нижним пределом частоты 200 Гц. Такой частотный порог выбирается на основе типичного диапазона птичьих звуков, поскольку частоты ниже 200 Гц обычно состоят из низкочастотных шумов (таких как шаги или ветер), которые не предоставляют полезной информации для задачи распознавания. Следовательно, Mel-спектрограммы служат информативным входным сигналом для нейронной сети.

Оценка функциональности модели

Для оценки качества работы модели применяются стандартные метрики классификации: precision (точность), recall (полнота) и f1-score (гар-

моническое среднее точности и полноты). Для каждого вида птиц формируется подробный отчет, содержащий значения указанных метрик по каждому классу. Дополнительно составляется сводная таблица, отражающая распределение предсказанных вероятностей принадлежности к различным видам, что позволяет проанализировать уверенность модели в своих предсказаниях [7].

Для интерпретации выходных значений модели используется фиксированный порог вероятности, равный 0.8: предсказание считается положительным для конкретного класса, если соответствующее значение на выходе модели превышает данный порог. Такой выбор обоснован стремлением минимизировать количество ложноположительных классификаций и повысить достоверность распознавания.

В табл. 1 представлен фрагмент обученной модели и ее результаты для 10 наиболее качественно распознанных видов птиц из 116. Таблица метрик качества распознавания видов птиц содержит следующие поля: precision, recall, f1-score и support. Последнее поле представляет количество экземпляров данных в тестовой выборке, для которой рассчитывались метрики. Представлены метрики как для каждого вида птиц по отдельности, так и усредненные значения по всем видам с помощью разных алгоритмов подсчета. В результате усредненные по всем классам (видам птиц) значения метрик качества предсказания, полученные для итоговой модели, составили: *precision* – 0.92, *recall* – 0.74, *f1-score* – 0.82.

Таблица 1

Результаты тренировки модели

Class of a bird	Precision	Recall	f1-score	Support
Caprimulguseuropaeus	1.00	0.99	0.99	158
Bombycillagarrulus	1.00	0.94	0.97	63
Laniusminor	1.00	0.92	0.96	90
Aegithaloscaudatus	0.98	0.95	0.96	97
Loxiapytyopsittacus	0.97	0.95	0.96	78
Haliaeetusalbicilla	0.98	0.93	0.95	98
Aegoliusfunereus	0.94	0.94	0.94	65
Accipiternisus	0.92	0.95	0.93	91
Clangaclanga	1.00	0.87	0.93	54
Coloeusmonedula	1.00	0.85	0.92	61
micro avg	*0.92*	*0.73*	*0.81*	*10 609*
macro avg	*0.92*	*0.74*	*0.82*	*10 609*
weighted avg	*0.91*	*0.73*	*0.80*	*10 609*
samples avg	*0.78*	*0.76*	*0.76*	*10 609*

Повышение точности модели распознавания голосовых сигналов птиц

Для обеспечения корректной работы модели распознавания и повышения точности оценки состояния биоразнообразия лесных экосистем Республики Беларусь были использованы обновленные версии инструментальных средств и алгоритмов модели. Это позволило применять новые усовершенствованные функции в архитектуре распознавания звуковых сигналов животных, таких как *EfficientNetV2B3* и оптимизатор *AdamW*, что способствовало оптимизации вычислений на GPU, а также обеспечило быструю сходимость к оптимальному решению по времени и количеству циклов обучения модели распознавания. В итоге повысились точность и скорость распознавания голосовых сигналов животных. Таким образом, модель была переобучена с использованием этих усовершенствований. Результаты данного эксперимента отражены в табл. 2. Увеличение параметров precision, recall и f1-score показывает, что повысилась точность предсказаний распознанных видов птиц.

Таблица 2

Результаты переобучения модели на обновленных версиях вычислительной среды

Параметр	Precision	Recall	f1-score	Support
micro avg	0,93	0,73	0,82	10 609
macro avg	0,94	0,75	0,83	10 609
weighted avg	0,93	0,73	0,81	10 609
samples avg	0,78	0,76	0,76	10 609

После обновления используемой вычислительной среды стало доступно семейство архитектур *efficientnet_v2*, которое является модернизированной версией архитектур EfficientNet первой версии. Были проведены два эксперимента на двух разных архитектурах *efficientnet_v2* с целью улучшения качества распознавания модели. Для первого эксперимента была взята архитектура модели распознавания EfficientNetV2B3 с параметром negative_slope для функции активации ReLU, равным 0,1. Результаты эксперимента представлены в табл. 3.

Таблица 3

Результаты эксперимента с архитектурой EfficientNetV2B3 с параметром negative_slope для функции ReLU, равным 0,1

Параметр	Precision	Recall	f1-score	Support
micro avg	0,92	0,73	0,82	10 609
macro avg	0,94	0,75	0,82	10 609
weighted avg	0,93	0,73	0,81	10 609
samples avg	0,78	0,76	0,76	10 609

Примеры сравнения изменения метрики precision в зависимости от порога вероятности модели с архитектурой EfficientNetV2B3 и параметром negative_slope для функции активации ReLU, равным 0,1, и модели, используемой в данный момент для видов птиц (певчий дрозд, обыкновенный снегирь), представлены на рис. 4.

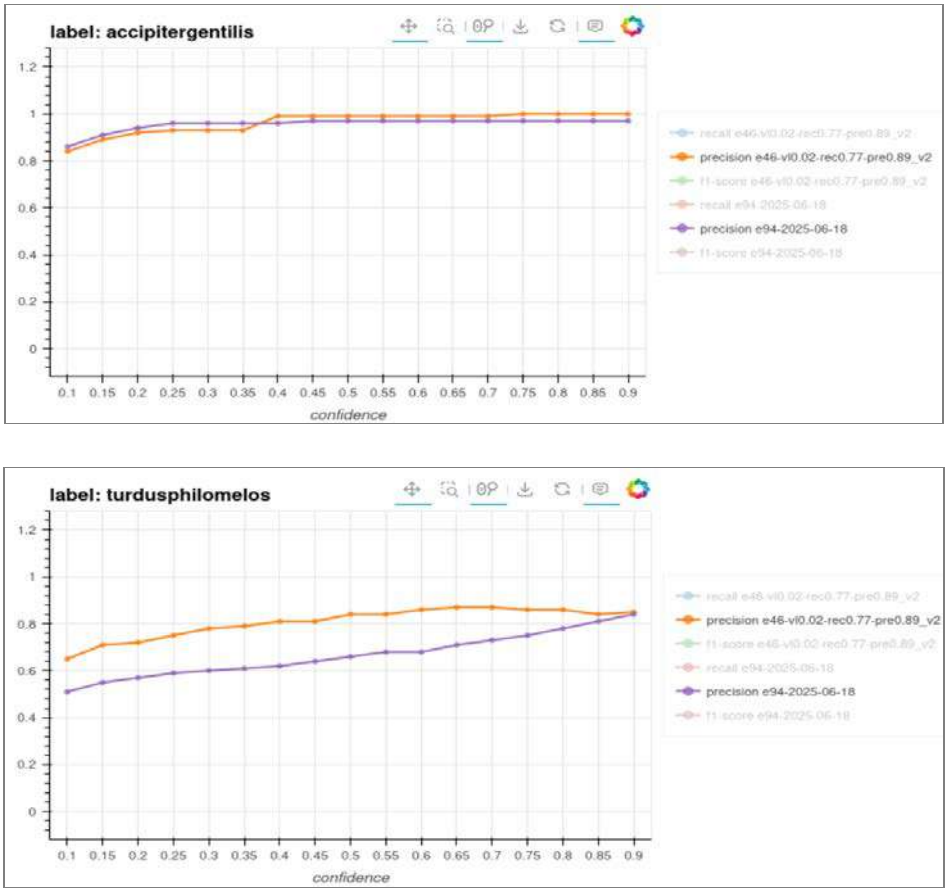


Рис. 4. Сравнение метрики precision модели с архитектурой EfficientNetV2B3 и параметром negative_slope для обыкновенного снегиря (слева) и певчего дрозда (справа)

Второй эксперимент проводился с использованием архитектуры модели распознавания EfficientNetV2M. Результаты эксперимента представлены в табл. 4. Значения результирующих метрик последней полученной модели на EfficientNetV2M оказались ниже, чем значения метрик модели на EfficientNetV2B3. Таким образом, модель распознавания с архитектурой EfficientNetV2B3 повысила точность оценки состояния биоразнообразия лесных экосистем с использованием экспериментального образца программного обеспечения автоматизированного распознавания голосовых сигналов животных в собранных базах данных.

Таблица 4

Результаты эксперимента с архитектурой модели EfficientNetV2M

Параметр	Precision	Recall	f1-score	Support
micro avg	0,86	0,47	0,61	10 609
macro avg	0,88	0,45	0,55	10 609
weighted avg	0,86	0,47	0,57	10 609
samples avg	0,51	0,49	0,49	10 609

Для пополнения базы данных новыми записями вокализаций птиц был использован датасет *WABAD*, который находится в открытом доступе [8]. Отобранные аудиозаписи *WABAD* были добавлены к имеющимся вокализациям птиц в основной базе данных. В результате был получен датасет, состоящий из 89 033 окон по всем предсказываемым видам птиц. На нем была обучена модель архитектуры EfficientNetB3. Результаты обучения представлены в табл. 5.

Таблица 5

Результаты обучения модели с использованием датасета WABAD

Параметр	Precision	Recall	f1-score	Support
micro avg	0,89	0,61	0,72	29 745
macro avg	0,92	0,63	0,74	29 745
weighted avg	0,89	0,61	0,71	29 745
samples avg	0,73	0,65	0,67	29 745

Результаты эксперимента показали, что по сравнению с моделью, обученной на наборе данных, который используется для обучения предыдущих моделей, полученная модель распознавания показывает значения метрик ниже. Одной из причин ухудшения результатов могут быть ошибки в аннотациях, так как датасет WABAD был взят из отрытых источников с готовыми аннотациями данных, которые не подвергались экспертной проверке на корректность. Эти ошибки могли оказать отрицательное влияние на обучаемость модели. Вторым отрицательным фактором может быть присутствие большого объема данных о видах птиц, не типичных для Республики Беларусь. В условиях ограниченного объема данных, близкого к минимально необходимому для качественного обучения, подобные вокализации могут оказывать существенное влияние на точность модели распознавания.

Для улучшения качества распознавания голосовых сигналов животных был проведен эксперимент с фильтрацией аннотаций. После шумоподавления часть вокализаций птиц становится слишком тихой и фактически недоступной для распознавания. Тем не менее они остаются в данных для обучения, поэтому соответствующие фрагменты аудио по-прежнему попадают в набор для обучения. Чтобы избежать включения таких «пустых» примеров, в цепочку предварительной обработки был добавлен этап фильтрации аннотаций. На этом этапе проверяются записи после удаления шумов и аннотации, которые больше не слышны, исключаются из датасета. Модель была обучена на нескольких наборах данных. Итоговые результаты отражены в табл. 6 и 7.

Таблица 6

Результаты обучения модели на очищенном наборе данных, включающем датасет WABAD

Параметр	Precision	Recall	f1-score	Support
micro avg	0,90	0,57	0,70	27 448
macro avg	0,93	0,60	0,71	27 448
weighted avg	0,90	0,57	0,69	27 448
samples avg	0,72	0,62	0,65	27 448

Таблица 7

Результаты обучения модели на очищенном наборе данных, включенных в общую базу данных

Параметр	Precision	Recall	f1-score	Support
micro avg	0,93	0,77	0,84	14 505
macro avg	0,94	0,77	0,84	14 505
weighted avg	0,93	0,77	0,83	14 505
samples avg	0,82	0,79	0,79	14 505

Модель, обученная на очищенном наборе данных без аудиозаписей датасета WABAD, показала лучшие значения метрик. Далее будет приниматься решение о том, какая модель будет использоваться в дальнейшем для обработки звуковых сигналов животных с целью повышения точности оценки состояния биоразнообразия лесных экосистем с использованием экспериментального образца программного обеспечения автоматизированного распознавания голосовых сигналов животных в собранных базах данных (в части технической реализации).

Заключение

Сочетание современных методов глубокого обучения и биоакустического анализа позволило создать эффективную автоматическую систему распознавания птиц. Архитектура EfficientNetB3 продемонстрировала высокую точность при умеренных вычислительных затратах, успешно преобразуя вокализации в Mel-спектрограммы и обеспечивая надежную работу даже на зашумленных записях. Интеграция современных методов ИИ, в частности предобученной сети EfficientNetB3, усовершенствованных алгоритмов обучения, оптимизатора AdamW и продуманной цепочки предобработки аудиоданных, позволяет достичь высокой точности автоматического распознавания голосовых сигналов птиц белорусской фауны. Полученные средние метрики (precision = 0.92, recall = 0.74, F1-score = 0.82) подтверждают конкурентоспособность модели по сравнению с традиционными методами мониторинга, которые требуют значительных временных и человеческих ресурсов.

Переход к архитектурам EfficientNetV2 дополнительно улучшил эффективность обучения: ускоренная сходимость и более эффективное использование GPU привели к росту всех ключевых показателей (precision, recall, f1-score) без увеличения вычислительной нагрузки.

Таким образом, разработанная система представляет собой практический инструмент для масштабируемого, непрерывного мониторинга биоразнообразия лесных экосистем, способного поддерживать охранные мероприятия и научные исследования, минимизируя человеческий фактор и повышая надежность данных. Полученные результаты демонстрируют практический потенциал интеллектуальных систем мониторинга птиц в охране природы и экологических исследованиях. Описанный подход упрощает идентификацию, создает инфраструктуру и открывает возможности применения в более сложных природных условиях. В дальнейшем планируется доработка модели, расширение базы данных и адаптация метода к новым задачам биоакустики, что позволит еще больше расширить возможности автоматических систем мониторинга дикой природы.

Список использованных источников

1. Sharma, S. A methodological literature review of acoustic wildlife monitoring using artificial intelligence tools and techniques / S. Sharma, K. Sato, B. P. Gautam // Sustainability. – 2023. – Vol. 15, no. 9. – P. 7128. – DOI: 10.3390/su15097128.

2. Insights and approaches using deep learning to classify wildlife / Z. Miao, K. M. Gaynor, J. Wang [et al.] // Scientific Reports. – 2019. – Vol. 9, no. 1. – URL: https://www.nature.com/articles/s41598-019-44565-w (date of access: 13.02.2021).

3. Stowell, D. Computational bioacoustics with deep learning: a review and roadmap / D. Stowel // Peer J. – 2022. – Vol. 10, art. e13152. – P. 1–46.

4. Аўтаматызацыя аналізу галасавых сігналаў птушак / Ю. С. Гецэвіч, Я. С. Зяноўка, А. А. Бакуновіч [і інш.] // Информатика. – 2024. – № 4. – С. 58–71.

5. The Belarusian list of bird species approved by the Belarusian ornitho-faunistic commission for 2021 and 2022 / N. V. Karlionova, A. V. Borodin, I. E. Samusenko, M. Y. Nikiforov // Зоологические чтения : сб. науч. ст., посвящ. 125-лет. д-ра биол. наук Ивана Николаевича Сержанина / ГрГУ им. Янки Купалы ; редкол.: О. В. Янчуревич (гл. ред.), А. В. Рыжая. – Гродно : ГрГУ, 2023. – С. 113.

6. Мадэль баз даных для тэхналогіі аўтаматызаванага распазнавання галасавых сігналаў жывёл / С. А. Гайдураў, Д. І. Латышэвіч, А. А. Бакуновіч [і інш.] // Развитие информатизации и государственной системы научно-технической информации (РИНТИ-2022) : докл. XXI Междунар. науч.-техн. конф., Минск, 17 нояб. 2022 г. – Минск : ОИПИ НАН Беларуси, 2022. – С. 236–240.

7. Developing an AI-powered bird sound recognition system for monitoring avian biodiversity in Belarus / Y. Zianouka, V. Dydo, A. Dydo [et al.] // Open Semantic Technologies for Intelligent Systems. – 2025. – № 9. – P. 299–308.

8. Pérez-Granados, C. WABAD: A World Annotated Bird Acoustic Dataset for Passive Acoustic Monitoring / C. Pérez-Granados, E. Sebastián-González. – 2025. – URL: https://zenodo.org/records/14191524 (date of access: 19.06.2025).

УДК 338.22

Суверенный искусственный интеллект как основа цифрового будущего Беларуси

Е. С. Жданович, В. А. Макренкова, Е. А. Алексеева✉
Витебский государственный
технологический университет, Беларусь
E-mail: vmakar555@mail.ru

Введение

Широкое использование цифровых технологий и искусственного интеллекта (ИИ) в разных сферах жизни общества (промышленность, транспорт, медицина, образование, государственное управление, финансовый сектор) в условиях международной конкуренции за технологическое лидерство обусловливает необходимость поддержания технологического суверенитета как фактора сохранения самостоятельности и независимости государства и его граждан.

Технологический суверенитет – это способность страны самостоятельно разрабатывать, производить и контролировать ключевые технологии и инфраструктуру, необходимые для экономической независимости и национальной безопасности, что создает возможность для принятия и реализации самостоятельных решений в сфере развития страны.

Технологический суверенитет требует наличия таких ключевых компонентов, как:

– техническая база: собственные дата-центры, облачные платформы и мощные компьютеры (без этого нельзя запускать современные сервисы);

– контроль над данными: информация о гражданах, компаниях и государстве должна храниться и обрабатываться внутри страны по своим законам (это защищает от кражи данных и внешнего давления);

– наличие специалистов: в стране должно быть достаточно программистов, инженеров и аналитиков, которые разбираются в ИИ и цифровых технологиях;

– установление собственных правил: необходимы законы, которые регулируют использование технологий, защищают личные данные и обеспечивают безопасность.

Сегодня технологический суверенитет позволяет обеспечить безопасность страны: контролировать работу энергетики, транспорта, финансов,

чтобы их нельзя было взломать или отключить извне; гарантировать устойчивость экономики, снижая зависимость от иностранных компаний, которые могут уйти, поднять цены или прекратить поддержку своих программ; самостоятельно решать, как развиваться, какие технологии и как использовать в соответствии со своими ценностями и интересами.

ИИ является мощным инструментом, который может как создать возможности для страны стать технологически сильнее, так и, наоборот, позволить ей попасть в еще большую зависимость от других государств. Для Беларуси, которая хочет уверенно развиваться, создание своих цифровых решений и ИИ является важной задачей.

Цифровой суверенитет можно определить как способность страны самостоятельно разрабатывать, использовать и контролировать важнейшие цифровые технологии.

Значение ИИ для обеспечения технологического суверенитета и информационной безопасности

ИИ способен брать на себя огромный объем рутинной, однообразной работы, требующей ручного труда: обработку документов, проверку заявок, составление отчетов. Он может автоматизировать эти задачи, делая их быстрее и без ошибок, а люди смогут сосредоточиться на более творческих и сложных вопросах. Кроме того, ИИ умеет анализировать гигантские массивы информации, которые не под силу человеку. Для правительства ИИ может выявлять скрытые тенденции в экономике, точнее прогнозировать доходы и расходы бюджета, заранее предупреждать о возможных кризисах и даже эффективнее бороться с мошенничеством. Не менее важна его роль в безопасности. Киберпреступники постоянно изобретают новые способы атак, и ИИ способен постоянно мониторить интернет-трафик, мгновенно выявляя подозрительную активность и нейтрализуя угрозы до того, как они нанесут ущерб. Наконец, создавая собственные программы на основе ИИ – умных помощников, систем перевода или прогнозирования, страна перестает быть просто потребителем и сама начинает создавать цифровые продукты, отвечающие ее уникальным потребностям [1].

Однако широкое использование ИИ порождает дополнительные риски. Активное применение зарубежных платформ ИИ, таких как решения от Google, Microsoft или Amazon, создает серьезные угрозы:

1. Технологическая зависимость. Если вся работа государственных органов или критически важных предприятий построена на иностранном ИИ, то в любой момент, в силу политических санкций или иных ограничений, доступ к этим системам может быть заблокирован или ограничен, что парализует работу целых отраслей.

2. Утечка данных. Когда информация обрабатывается на чужих серверах, она автоматически попадает под юрисдикцию другой страны. Это означает, что персональные данные граждан, коммерческие секреты компаний и даже государственная тайна могут оказаться в распоряжении иностранных правительств или корпораций.

3. Несовершенство систем. ИИ, обученный преимущественно на данных из других стран, плохо понимает местные языковые нюансы, культурные особенности и специфику законодательства. В результате его прогнозы и рекомендации для Беларуси могут быть неточными или вовсе неверными, что приведет к ошибочным управленческим решениям.

4. Увеличение технологического разрыва. Ведущие державы, понимая стратегическую ценность ИИ, могут сознательно ограничивать доступ к самым передовым алгоритмам и вычислительным мощностям, оставляя другим странам устаревшие версии и обрекая их на роль вечных догоняющих.

Именно эти риски и породили то, что сегодня в мире называют «гонкой за суверенный ИИ». Страны осознали, что обладание собственными мощными компьютерами для обработки данных и своими, национальными наборами информации для обучения алгоритмов из плоскости престижа или экономической выгоды переросло в фундаментальный вопрос национальной безопасности, условие для самостоятельного принятия решений и гарантию того, что государство сможет защитить своих граждан и свои интересы в цифровую эпоху. Это гонка, проиграть в которой означает добровольно отказаться от части своего суверенитета [2].

Республика Беларусь на пути к «суверенному ИИ»

Страны поняли, что иметь свои вычислительные мощности и свои данные для обучения ИИ – это вопрос выживания в 21 в.

Ответом Беларуси на эти технологические вызовы стало знаковое событие – создание в июне 2025 г. первой национальной платформы ИИ. Разработанная Национальным банком и Центром банковских технологий (ЦБТ) эта платформа стала практическим доказательством способности страны создавать сложные ИИ-системы собственными силами. Ее архитектура изначально проектировалась как полностью автономное решение, способное функционировать даже в условиях полного отсутствия доступа к зарубежным сервисам и облачным платформам. Это стратегическое решение основано на трех фундаментальных принципах: максимальной технологической независимости, безусловном приоритете безопасности и на адаптации под специфические потребности местных пользователей и отраслей [3].

Ключевыми компонентами платформы являются несколько взаимосвязанных модулей. Основу составляют продвинутые языковые модели, специально дообученные для эффективной работы с русским и белорусским языками. Эти модели не просто понимают запросы, а способны генерировать связные тексты, проводить смысловой анализ сложных документов, классифицировать информацию и поддерживать диалог в форме интеллектуальных чат-ботов, что критически важно для автоматизации клиентских сервисов и поддержки сотрудников. Второй критически важный элемент – инструменты роботизации процессов, которые берут на себя массовые, повторяющиеся операции. В банковской сфере, например, они автоматически проверяют пакеты документов на соответствие критериям, обрабатывают заявки на продукты и услуги, а также формируют регламентированные отчеты, что не только ускоряет процессы в разы, но и сводит к минимуму человеческий фактор и связанные с ним ошибки. Третья составляющая – это мощные аналитические инструменты, предназначенные для глубокого анализа огромных массивов как структурированных, так и неструктурированных данных. Они выявляют скрытые закономерности, строят прогнозные модели для оценки, например, кредитных рисков или рыночной динамики, а также в реальном времени обнаруживают подозрительные транзакции и сложные схемы мошенничества, которые часто остаются незамеченными при ручном контроле.

Вопросу безопасности в архитектуре платформы уделено первостепенное внимание. Самый радикальный и эффективный принцип – это работа в режиме полной изоляции от глобальной сети. Платформа физически не подключена к интернету, что создает непреодолимый барьер для внешних кибератак, несанкционированного доступа и удаленного хищения данных. Принцип data sovereignty реализован здесь в абсолютной форме: все данные, без исключения, обрабатываются и хранятся на серверных мощностях, расположенных исключительно на территории Беларуси. Это не только гарантирует их защиту в соответствии с национальным законодательством о персональных данных и коммерческой тайне, но и полностью выводит их из-под действия иностранных юрисдикций. При этом платформа обладает значительной гибкостью – ее модульная архитектура позволяет настраивать и масштабировать функционал под конкретные задачи самых разных секторов экономики – от финансов и государственного управления до здравоохранения, образования и промышленного производства.

Пилотной площадкой для внедрения платформы стал финансовый сектор. Здесь она уже доказала свою эффективность, решая такие задачи, как оперативная обработка обращений клиентов через чат-ботов, круглосуточный мониторинг транзакций для выявления мошеннических паттернов и автоматизированная быстрая оценка кредитоспособности заемщи-

ков. Успешный опыт в банковской сфере заложил основу для стратегии масштабирования. В ближайшей перспективе платформа готовится к внедрению в систему государственного управления для анализа социально-экономических данных и оптимизации оказания госуслуг, в промышленность – для прогнозирования спроса и предиктивного обслуживания оборудования, а также в медицину – для помощи в диагностике и анализа медицинских изображений. Таким образом, эта суверенная разработка постепенно превращается в универсальный технологический каркас для цифровой трансформации всей страны, обеспечивая ее технологическую независимость на долгосрочную перспективу [3].

Создание суверенной платформы искусственного интеллекта, безусловно, является важным достижением. Этот проект следует рассматривать как составную часть масштабной государственной стратегии, изложенной в Концепции обеспечения суверенитета Республики Беларусь в сфере цифрового развития до 2030 года (URL: https://pravo.by/document/?guid=3961&p0=C22401074).

Данный основополагающий документ, утвержденный постановлением Совета Министров Республики Беларусь в декабре 2024 г., задает системное видение того, как страна должна укреплять свою технологическую независимость в современном мире. Важно подчеркнуть, что концепция трактует цифровой суверенитет не как стремление к полной изоляции или технологическому национализму, а как стратегическую способность государства самостоятельно, без внешнего диктата, определять вектор своего цифрового развития, защищать интересы граждан и бизнеса в цифровой среде и сохранять устойчивость национальной экономики и инфраструктуры в условиях внешнего давления и санкционных ограничений [4].

Для достижения поставленной цели документ определяет несколько ключевых стратегических направлений. Во-первых, это планомерное развитие отечественного программного обеспечения через всестороннюю поддержку белорусских разработчиков и создание для них приоритетных условий для внедрения их продуктов в государственном секторе и на предприятиях, имеющих критическое значение для экономики. Во-вторых, концепция предполагает целенаправленное снижение импортозависимости в сфере программного обеспечения, предусматривая постепенный, но настойчивый процесс замены иностранного ПО, особенно в системах, связанных с национальной безопасностью, функционированием критической информационной инфраструктуры и обработкой персональных данных граждан. В-третьих, для снижения затрат и повышения контроля над исходным кодом государственная политика будет активно стимулировать использование свободного и открытого программного обеспечения (СПО/Open Source), что также способствует развитию собственных компетенций.

Кадровое обеспечение информационного суверенитета

Важным аспектом всей стратегии является подготовка собственных IT-специалистов, подразумевающая масштабные инвестиции в образование – от школьной скамьи до вузовских программ и курсов повышения квалификации, чтобы обеспечить страну устойчивым потоком высококвалифицированных кадров в сферах ИИ, анализа данных и кибербезопасности. Наконец, концепция закрепляет необходимость сохранения полного государственного контроля над ключевыми элементами цифровой инфраструктуры, такими как магистральные каналы связи и национальные датацентры, которые являются физической основой цифрового суверенитета.

План также включает в себя комплекс мер по стимулированию инноваций: он предусматривает поддержку фундаментальных и прикладных научных исследований в области ИИ, создание благоприятных условий для стартапов, работающих с передовыми цифровыми технологиями, и активное внедрение интеллектуальных систем в процессы государственного управления для повышения их эффективности, прозрачности и скорости принятия решений. Таким образом, Концепция до 2030 г. представляет собой не просто декларацию о намерениях, а всеобъемлющую дорожную карту, которая координирует усилия государства, бизнеса и научного сообщества для построения в Беларуси устойчивой и технологически независимой цифровой экосистемы.

Несмотря на очевидные успехи, путь Беларуси к полноценному цифровому суверенитету сопряжен с комплексом серьезных вызовов, требующих системного подхода и последовательных усилий. Ключевой проблемой остается глубокая структурная зависимость от иностранного программного обеспечения: множество государственных учреждений и предприятий критической инфраструктуры десятилетиями работают на зарубежных технологических стеках, и их миграция на отечественные аналоги сопряжена не только с колоссальными финансовыми затратами, но и с операционными рисками, необходимостью переобучения персонала и обеспечением бесперебойной совместимости систем. Эта проблема усугубляется тем, что в ряде критически важных областей, таких как операционные системы корпоративного уровня, системы автоматизированного проектирования (CAD/CAM/CAE) или сложные системы управления базами данных, конкурентоспособных белорусских аналогов попросту не существует, что создает технологические вакуумы, которые невозможно быстро заполнить. Еще одним структурным вызовом является экспортная ориентация национального IT-сектора: значительная часть талантливых разработчиков и перспективных стартапов сфокусирована на выполнении заказов для иностранных рынков, что обусловлено их большей финансо-

вой привлекательностью и меньшей бюрократической нагрузкой по сравнению с внутренними проектами. Создание конкурентоспособных решений в области искусственного интеллекта упирается в проблему ресурсов – оно требует не только многомиллионных инвестиций в дорогостоящее вычислительное оборудование (например, кластеры графических процессоров), но и острой конкуренции за узкопрофильных специалистов мирового уровня, на которых сегодня идет настоящая «охота». Наконец, нормативно-правовая база отстает от темпов технологического развития: отсутствуют четкие законы, регулирующие вопросы ответственности за решения, принимаемые алгоритмами; этические стандарты применения ИИ и механизмы сертификации таких систем, что создает правовую неопределенность и сдерживает их широкое внедрение в чувствительных сферах.

Для преодоления данных барьеров необходим стратегический и скоординированный ответ. Первоочередной задачей является развитие и масштабирование существующих отечественных разработок, таких как платформа ИИ ЦБТ, путем создания вокруг нее открытой экосистемы, которая позволит сторонним разработчикам и отраслевым компаниям строить специализированные приложения, тем самым значительно расширяя ее функциональность и сферы применения. Параллельно необходимо активнее и стратегически внедрять модели использования свободного и открытого программного обеспечения, что позволяет, не изобретая велосипед заново, адаптировать проверенные мировые наработки под национальные требования, экономя время и ресурсы. Фундаментальным условием долгосрочного успеха является тотальная работа с кадрами – интеграция основ программирования, анализа данных и кибербезопасности в школьные программы, модернизация университетских курсов, открытие специализированных магистратур и создание программ по привлечению в страну ведущих международных экспертов для обмена опытом. Мощным катализатором процесса должно выступить само государство, выступая в роли «первого клиента» для белорусских ИИ-решений через их пилотное внедрение в процессы оказания госуслуг, работы органов власти и государственных закупок, формируя тем самым стабильный внутренний спрос. Учитывая глобальный характер технологий, важно выстраивать стратегическое партнерство с дружественными странами, которые также движутся в направлении технологического суверенитета, для совместных научных исследований, согласования стандартов и кооперации в создании сложных технологических продуктов, что позволяет объединить ресурсы и компетенции для достижения общих целей.

Заключение

Сегодня технологии и ИИ – это уже не просто удобные инструменты, а вопрос независимости и безопасности для любой страны. Для Беларуси, которая находится в непростых международных условиях, создание своих технологий стало не просто желанием, а настоятельной потребностью. От этого зависит наша способность самостоятельно управлять экономикой, защищать данные граждан и обеспечивать работу важных систем – от банков до энергетики.

Понимание данной угрозы заставило страну действовать. Первым серьезным достижением стало создание в 2025 г. собственной платформы ИИ. Этот проект доказал, что у Беларуси есть не только желание, но и возможности делать сложные и безопасные цифровые системы своими силами, даже когда доступ к зарубежным технологиям ограничен. Полученный успех был закреплен принятием государственного плана – Концепции цифрового суверенитета до 2030 года. Этот документ стал подробной инструкцией, как достичь технологической независимости: развивать свое программное обеспечение, готовить своих специалистов, строить свою инфраструктуру и принимать нужные законы.

Однако самый сложный этап еще впереди. Нас ждут серьезные трудности. Многие госучреждения и компании до сих пор работают на иностранном программном обеспечении, и перейти на свое – дорого и сложно. Нужно не только закупать технику, но и переучивать людей, переносить огромные объемы данных. Другая большая проблема – кадры. Следует не только готовить больше программистов, но и создавать для них интересные проекты и хорошие условия внутри страны, чтобы их таланты работали на развитие Беларуси, а не уходили за границу. Кроме того, в некоторых областях, например сложных системах проектирования, своих аналогов пока нет, а создание современных систем ИИ требует огромных денег на мощные компьютеры. Успех всего этого большого дела зависит от одного: смогут ли государство, бизнес и ученые работать вместе, как одна команда. Если они объединят свои силы, деньги и знания, то Беларусь сможет не только защитить себя от внешних угроз, но и стать полноценным участником мирового технологического рынка, предлагая свои надежные и безопасные решения. В конечном счете, именно в этом единстве, четком плане и упорной работе заключается основа для будущего процветания и безопасности нашей страны в цифровую эпоху.

Список использованных источников

1. Гонка за суверенный ИИ – больше, чем независимость // VIRTRE. – 2022. – URL: https://virtre.ru/articles/artificial-intelligence/gonka-za-suverennyj-ii-bolshe-chem-nezavisimost (дата обращения: 06.10.2025).

2. Цифровой суверенитет: вызовы и перспективы // Форум IT-образования ETU. – 2022. – Т. 5. – URL: https://discourse.etu.ru/ru/arhiv/2022/5/81-90 (дата обращения: 06.10.2025).

3. Пресс-релиз: Суверенная платформа искусственного интеллекта // Центр Банковских Технологий – ЦБТ. – 2025. – URL: https://www.cbt.by/press-reliz-suverennaya-platforma-iskusstvennogo-intellekta/ (дата обращения: 06.10.2025).

4. Концепция обеспечения «цифрового суверенитета» Беларуси // Портал MYFIN.BY. – 2025. – URL: https://myfin.by/article/tekhnologii/predstavlena-koncepcia-obespecenia-cifrovogo-suvereniteta-belarusi-cto-v-dokumente-34898 (дата обращения: 06.10.2025).

УДК 004.89:37.01

Архитектура адаптивной образовательной среды на основе гибридных баз знаний и генеративных моделей

М. П. Исаев
ООО «Девиар ЭйАй», Минск, Беларусь
E-mail: admin@dewiar.com

Введение

Стремительное развитие больших языковых моделей (LLM) открывает беспрецедентные возможности для трансформации образовательного процесса. Перспектива создания персонализированных интеллектуальных тьюторов, способных адаптироваться к темпу и потребностям каждого учащегося, является одной из наиболее значимых целей современной педагогической науки. Однако существующие подходы к интеграции LLM в образование выявили ряд системных проблем.

Распространенная практика дообучения (fine-tuning) моделей на узкоспециализированных данных сопряжена с фундаментальными недостатками. Ключевыми из них являются риск «катастрофического забывания» [2, 8], когда модель теряет ранее усвоенные базовые способности, и проблема генерации фактически неверной информации («галлюцинаций») [9, 10]. В образовательной среде, где точность и верифицируемость знаний имеют первостепенное значение, такие риски недопустимы.

Целью настоящей работы является разработка архитектуры, которая позволяет использовать сильные стороны современных LLM (рассуждение, генерация текста), нивелируя их недостатки. В работе предлагается сместить акцент с модификации самой модели на создание сложной, многоуровневой архитектуры управления знаниями на основе принципа Retrieval-Augmented Generation (RAG) [1].

Предлагаемая архитектура системы

В основе предлагаемого решения лежит принцип RAG, согласно которому большая языковая модель используется не как носитель знаний, а как универсальный обработчик информации, которую ей предоставляет внешняя, верифицированная база знаний. Весь технологический цикл разбит на два ключевых этапа: формирование «живой библиотеки» (базы знаний) и интеллектуальный механизм извлечения данных для генерации ответа.

Формирование мультимодальной базы знаний

Ключевой задачей является преобразование утвержденных учебных материалов в структурированный, машиночитаемый формат. В отличие от простого индексирования текстов предлагаемый подход учитывает мультимодальную природу учебников.

Обработка текста и таблиц. Текстовые данные сегментируются на осмысленные фрагменты (чанки). Табличные данные подвергаются сериализации: большие таблицы разбиваются на логические единицы (строки или группы строк) с обязательным сохранением контекста заголовков для каждого фрагмента.

Обработка визуальных данных. Графики, схемы и диаграммы обрабатываются с помощью моделей компьютерного зрения (например, на базе архитектуры Vision Transformer [6]) и специализированных алгоритмов chart understanding [11], которые извлекают структурированные описания визуальных элементов. Эти описания затем векторизуются, позволяя осуществлять семантический поиск по визуальному контенту.

Контекстное обогащение и построение графа знаний. Для каждого фрагмента данных LLM используется для генерации метаданных: краткого резюме о его месте в общей структуре книги, выделения ключевых идей и сущностей. Эти данные формируют граф знаний, отражающий семантические связи между различными концепциями, что является основой для более глубокого понимания контекста [7].

Гибридный механизм извлечения информации

Для поиска наиболее релевантной информации по запросу учащегося используется гибридная система, объединяющая сильные стороны нескольких подходов.

Семантический (векторный) поиск. Для нахождения фрагментов, близких по смыслу к запросу, используются современные мультиязычные модели представления предложений, такие как BGE-M3 [3] и Qwen Embedding [4]. Эти модели обеспечивают высокую точность поиска за счет более плотных и универсальных эмбеддингов, обученных на многоязычных корпусах и мультимодальных данных.

Полнотекстовый поиск. Для поиска по ключевым словам, терминам и формулам применяются встроенные механизмы полнотекстового поиска PostgreSQL (tsvector/tsquery) и расширение pgvector, позволяющее объединять семантический и текстовый поиск в единой среде хранения данных [5].

Графовый поиск. Запросы к графу знаний позволяют извлекать связанные понятия и темы, что обогащает контекст и обеспечивает более глубокое понимание.

Результаты всех поисковых систем ранжируются и объединяются с помощью алгоритма Reciprocal Rank Fusion (RRF) [6], доказавшего свою эффективность при объединении разнородных источников.

Обсуждение и перспективы

Предложенная архитектура обеспечивает ряд ключевых преимуществ перед методом дообучения:

Верифицируемость. Каждый ответ системы может быть напрямую соотнесен с первоисточником (учебником, страницей, параграфом), что существенно снижает риск «галлюцинаций» [9, 10], но не исключает необходимость дополнительной автоматической проверки фактов.

Масштабируемость. Обновление учебной программы требует лишь добавления новых материалов в базу знаний, а не дорогостоящего и длительного переобучения LLM.

Безопасность. Система может быть полностью развернута в закрытом локальном контуре, что гарантирует контроль над данными и соблюдение нормативных требований.

Потенциал данной архитектуры не ограничивается созданием виртуальных репетиторов. На ее основе могут быть реализованы ассистенты для преподавателей (генерация учебных планов и заданий), а также интерактивные среды, где ИИ выступает в роли консультанта, оперирующего исключительно верифицированными материалами.

Заключение

Разработанная архитектура представляет собой комплексное решение для создания эффективных и безопасных интеллектуальных образовательных систем. Сочетание принципа RAG с продвинутыми методами обработки мультимодальных данных и гибридного поиска позволяет создавать универсальные и адаптивные системы управления знаниями, устойчивые к изменениям и полностью контролируемые. Будущее ИИ в образовании, по мнению автора, лежит в области создания совершенных архитектур знаний, а не в бесконтрольном усложнении моделей.

Список использованных источников

1. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / P. Lewis, E. Perez, A. Piktus [et al.] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 9459–9474.

2. McCloskey, M. Catastrophic interference in connectionist networks: The sequential learning problem / M. McCloskey, N. J. Cohen // Psychology of Learning and Motivation. – 1989. – Vol. 24. – P. 109–165.

3. BGE M3: A Unified Multilingual, Multimodal, Multi-vector Embedding Model / J. Chen, S. Xiao, P. Zhang [et al.] // arXiv preprint. – 2024. – URL: https://arxiv.org/pdf/2402.03216v3 (date of access: 24.09.2025).

4. Qwen3 Embedding. – URL: https://github.com/QwenLM/Qwen3-Embedding (date of access: 24.09.2025).

5. PostgreSQL Global Development Group. PostgreSQL 16 Documentation: Full Text Search and pgvector Extension. – 2024. – URL: https://www.postgresql.org/docs/current/textsearch.html (date of access: 24.09.2025).

6. Cormack, G. V. Reciprocal Rank Fusion Outperforms Condorcet and Individual Rank Learning Methods / G. V. Cormack, C. L. Clarke, S. Buettcher // SIGIR Conf. on Research and Development in Information Retrieval, Boston, MA, USA, 19–23 July 2009. – Boston, 2009. – P. 758–759.

7. An Image is Worth 16x16 Words: Transformers for Image Recognition at Scale / A. Dosovitskiy, L. Beyer, A. Kolesnikov [et al.] // arXiv preprint. – 2020. – URL: https://arxiv.org/abs/2010.11929 (date of access: 24.09.2025).

8. A Survey on Knowledge Graphs: Representation, Acquisition, and Applications / S. Ji, S. Pan, E. Cambria [et al.] // IEEE Transactions on Neural Networks and Learning Systems. – 2021. – Vol. 33, iss. 2. – P. 494–514.

9. Overcoming Catastrophic Forgetting in Neural Networks / J. Kirkpatrick, R. Pascanu, N. Rabinowitz, J. Veness // Proceedings of the National Academy of Sciences (PNAS). – 2017. – Vol. 114, no. 13. – P. 3521–3526.

10. On Faithfulness and Factuality in Abstractive Summarization / J. Maynez, S. Narayan, B. Bohnet, R. McDonald // Proc. of the 58th Annual Meeting of the Association for Computational Linguistics (ACL), Online, July 2020. – Association for Computational Linguistics, 2020. – P. 1906–1919.

11. A Survey on Hallucination in Large Language Models / L. Huang, W. Yu, W. Yu [et al.] // arXiv preprint. – 2023. – URL: https://arxiv.org/abs/2311.05232 (date of access: 24.09.2025).

12. ChartOCR: Data Extraction from Charts via Vision-Language Pre-training / J. Luo, Z. Li, J. Wang, C.-Y. Lin // Proc. of the IEEE/CVF Conf. on Computer Vision and Pattern Recognition (CVPR). – 2023. – URL: https://www.microsoft.com/en-us/research/wp-content/uploads/2020/12/WACV_2021_ChartOCR.pdf (date of access: 24.09.2025).

УДК 004.8

Разработка электронной базы данных Автопедии, интерактивных методов подготовки оценщиков, поврежденных дорожных транспортных средств и возможности использования искусственного интеллекта

В. В. Капустин
Общественное объединение «Белорусская ассоциация экспертов и сюрвейеров на транспорте» (ОО «БАЭС»),
Минск
E-mail: baes_ocenka@mail.ru

Автомобилизация в Республике Беларусь

В настоящее время развитие автомобилестроения приводит к расширению сферы использования автомобиля благодаря постоянному развитию его конструкции и безопасности, эффективности и экономичности, ряда других качеств. Эта отрасль расширяется и затрагивает другие сферы деятельности человека. В последнее время применяется новое определение «автомобильность» и ее концепция. Оно образовалось на стыке нескольких научных областей знаний: технических наук и экономики дорожных транспортных средств (далее – ТС), маркетинга, культурологии и дизайна, городского планирования и градостроительства, социологии, медицины и физиологии человека.

Автомобиль является важнейшим объектом индивидуального потребления, который обеспечивает своему владельцу (пользователю) социальный статус через знаковые ценности (скорость, безопасность, карьерный успех, свободу передвижения, мужественность, соревнование). Вокруг автомобиля образовалась культура, в рамках которой формируются литературные и художественные образы, символы.

Таким образом, автомобиль стал занимать важное место в жизнедеятельности человека и качестве его жизни, обеспечивая комфорт, скорость и удобство при передвижении.

Электронная база данных Автопедии

В процессе развития автомобильности важное место занимают квалификация и система подготовки и повышения квалификации специалистов по автотранспорту, информационное и программное обеспечение.

В соответствии с нормативными документами, программами цифрования и импортозамещения в Республики Беларусь сотрудниками и членами ассоциации разработана электронная версия комплекса «Автопедия». Комплекс включает информационное и программное обеспечение по автомобильной тематике. База данных используется учреждением образования «Центр повышения квалификации руководящих работников и специалистов финансовой системы» (далее – Центр) для подготовки оценщиков ТС и другого поврежденного в ДТП имущества. Электронная версия Автопедии размещается в сети интернет как платформа сайта ассоциации ОО «БАЭС» www.autoexp.org. База данных используется в работе оценщиками ТС, специалистами автотранспорта, экспертами. Автопедия может быть рекомендована как пособие по профилю автомобильности, студентам, учащимся и автомобилистам.

Автопедия включает основные разделы:

– учебные пособия объемом 4478 страниц текста и графики по устройству автомобилей с двигателем внутреннего сгорания и электромобилей, диагностике, ремонту и возможной их утилизации, глоссарий;

– 136 учебно-методических фильмов ведущих организаций автосервиса по соответствующим темам Автопедии (активные ссылки по тексту тем);

– программы подготовки и повышения квалификации специалистов по оценке поврежденного имущества;

– методику интерактивного образования по программе подготовки;

– информацию об использовании разработанного по импортозамещению Объединенного программного комплекса (далее – ОПК+) для идентификации ТС, расчету стоимости восстановительного ремонта после ДТП и размера причиненного ущерба.

Разделы Автопедии по теме автомобилизации приведены на рис. 1. Функциональная схема электронной версии Автопедии доступна и удобна для пользователя. Она разработана в виде оптимальной иерархической системы доступа по 550 темам. Это позволит пользователям в удаленном режиме получать доступ с любого программного средства, изучать интересующую его тему или задавать конкретные вопросы в процессе работы, при обучении, самоподготовке.

Планируется также использовать некоторые функции обратной связи в режиме общедоступной универсальной интернет-энциклопедии, материалы которой (за исключением нормативных документов и методик) могут дополняться пользователями в корректной редакции. Каждый желающий в пределах предложенной тематики, касающейся устройства новых моделей автомобилей, тенденций развития, технологии изготовления автомобиля, его идентификации, осмотра, расчета стоимости, определения ущерба, использования методик и рекомендаций, может изложить свою пози-

цию по конкретному вопросу, получить консультацию. При необходимости специалисты ОО «БАЭС» и эксперты могут оказывать помощь – давать пояснения по конкретному вопросу и обращению.

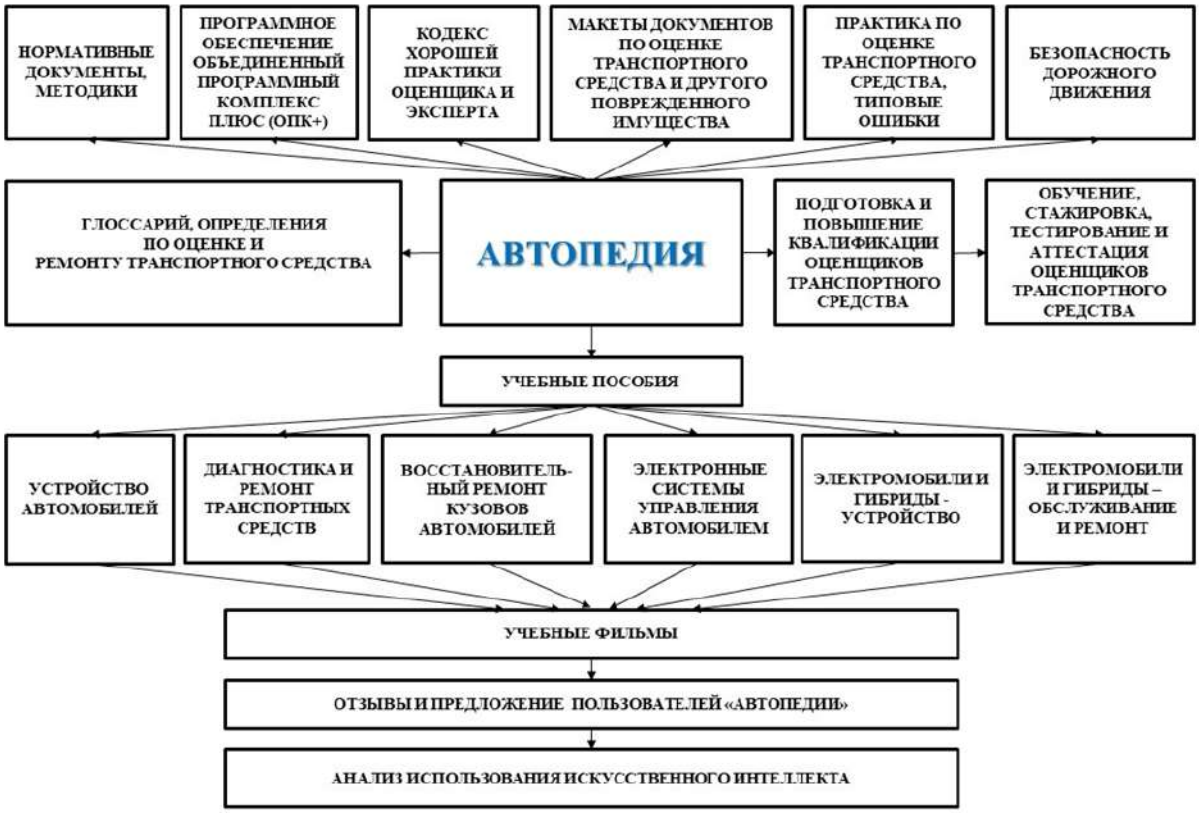


Рис. 1. Функциональная схема электронной версии Автопедии

В ассоциации ОО «БАЭС» создан технический Совет из разработчиков Автопедии и независимых экспертов. При поступлении конкретных предложений пользователей: новых разработок, статей, предложений – они будут рассмотрены и при их положительных оценках могут быть размещены в Автопедии в специальном разделе.

Автопедия также становится платформой для размещения юридическими и физическими лицами по автомобильной тематике.

Интерактивные методы образования

В Автопедии, а также на сайте ОО «БАЭС» размещена разработанная методика использования интерактивных методов по программе подготовки и повышения квалификации специалистов. Учитывается ряд особенностей. Они заключаются в следующем: у слушателей по подготовке обычно различный уровень знаний и опыта, сложившейся практики прежней работы, различные возрастные группы. Использование только известных методов,

как лекционных, так и практических занятий, требует дополнительных методов интерактивной формы с применением обратной связи и новых средств коммуникаций (рис. 2). Это повышает качество обучения и в то же время уменьшает его срок.

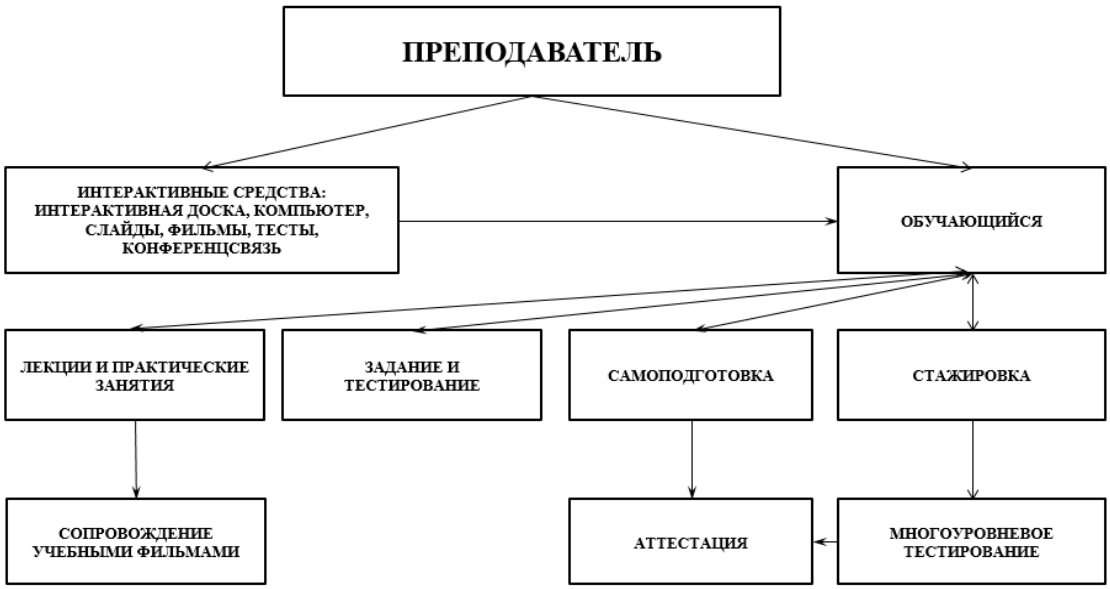


Рис. 2. Практическое применение интерактивного обучения

В учебном процессе дополнительно используются интерактивная доска, компьютеры для слушателей, интерактивные формы проведения занятий. Это тематические учебные фильмы, базы данных Автопедии, дидактические материала на занятиях, создание и передача слушателям электронной библиотеки по программе подготовки и содержание основных положений лекций и отдельных практических занятий.

Предусмотрена возможность сотрудничества преподавателя и слушателей, разбор и обсуждение конкретных ситуаций и практики оценщика. Для решения учебных задач программы подготовки используются следующие основные интерактивные формы: круглый стол (дискуссия, дебаты); мастер-класс специалиста-оценщика, эксперта; анализ конкретных экстремальных ситуаций с участием специалистов – работодателя, психолога; стажировка в организациях по оценке ТС; индивидуальная работа со слушателем опытного оценщика-эксперта (наставника) при прохождении стажировки.

Во время стажировки слушатель работает вместе с наставником-экспертом: проводит идентификацию поврежденного реального ТС, устанавливает сложность ремонтных воздействий, производит расчет стоимости восстановительного ремонта и другие процедуры, оформление документов.

Он также ведет дневник и по завершении стажировки составляет отчет, защищает его перед тестированием. Отчет хранится в архиве Центра.

Интерактивные методы обучения позволяют слушателям быть активными участниками образовательного процесса. Это развивает умение работать в команде, лидерские качества при обсуждении, дискуссиях или принятии решений (например, при различных интересах сторон: потерпевшего в ДТП и виновника).

Преподаватель становится не только транслятором знаний, но и наставником, обеспечивающим успешную групповую коммуникацию. Работа с группой слушателей во время выполнения программы подготовки перестраивается от монологической к диалогической.

Занятия, проводимые с использованием интерактивных методов, могут составлять не менее 20 % от лекционных и практических занятий.

В период сложившихся ситуаций используемая методика позволила организовать дистанционное обучение. В этом случае транслировалось проведение занятий на рабочее место присутствующего слушателя.

Подготовка слушателей по оценке ТС проводится по трем программам с соответствующим тестированием. Протокол тестирования хранится в учреждении образования, копию получает слушатель. В случае неуспешного тестирования слушатель имеет возможность пройти тестирование через месяц при успешной самоподготовке.

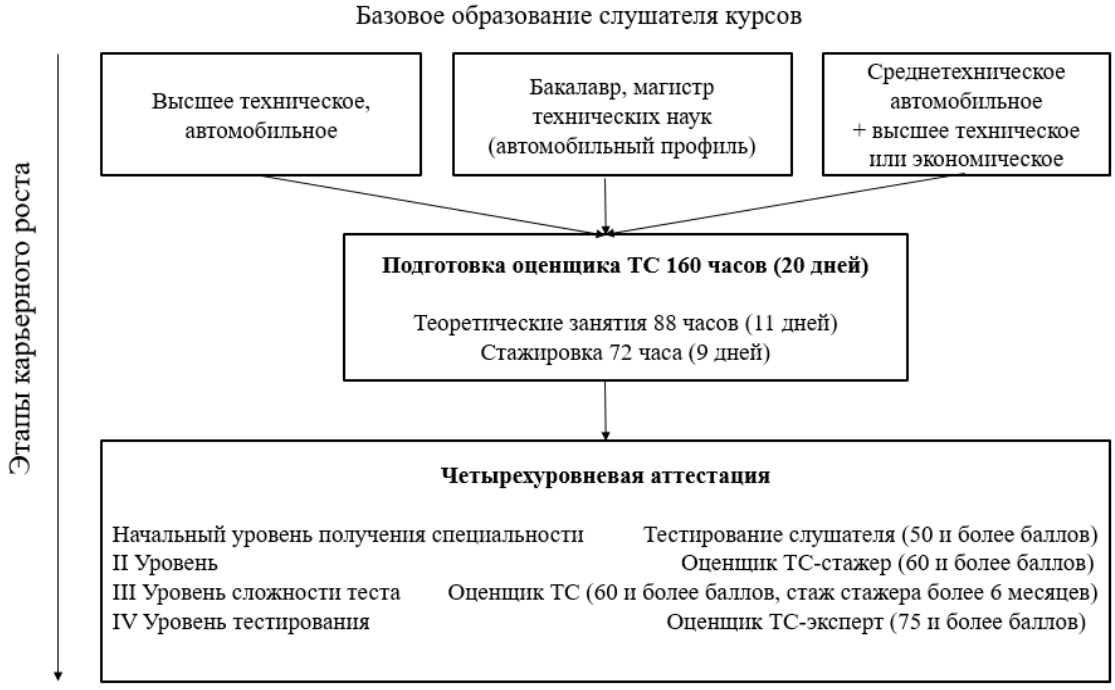


Рис. 3. Аттестация оценщика ТС

Образовательные программы включают: подготовку и повышение квалификации оценщика ТС с последующей четырехуровневой аттестацией его профессионального роста и повышение квалификации оценщика поврежденного иного имущества (рис. 3).

Возможность использования искусственного интеллекта

ОО «БАЭС» и ее предприятие могут предложить сотрудничество по следующим направлениям:

1. Электронная версия учебных пособий и база информационного обеспечения Автопедии для их использования в образовательных программах обучения учащихся и студентов автомобильного профиля, повышения квалификации специалистов с высшим инженерным образованием.

2. Использование специалистами разработанного в Республике Беларусь программного продукта ОПК+ для идентификации, расчетов остаточной стоимости ТС, их восстановительного ремонта или утилизации, не подлежащего санкциям.

3. При подготовке и повышении квалификации специалистов с высшим инженерным образованием использование апробированных интерактивных методов и организационного инжиниринга.

4. Планируется использование искусственного интеллекта.

4.1. При разработках алгоритмов анализа основных тенденций по тематике развития автомобилизации, анализу больших массивов переменных данных информационного и программного обеспечения Автопедии и других источников; при обосновании вариантов ответов на вопросы пользователей по автомобильной тематике с участием экспертов ОО «БАЭС».

4.2. Проведены пробные оценки перспективы развития автомобилей с двигателем внутреннего сгорания, электромобилей и автомобилей-гибридов в части изменения их стоимости в зависимости от времени эксплуатации. Для исследования применялись: нейросеть DeepSeek, генеративный чат-бот Grok 4-fast, а также поисковая система Perplexity на основе искусственного интеллекта.

4.3. Расширение возможностей интерактивного обучения с использованием искусственного интеллекта: персонализированное обучение – наставничество; генерация учебных материалов, презентаций, планов занятий и их адаптация, что сокращает время подготовки и повышает качество обучения; вовлечение слушателей во взаимодействие, мотивацию на их выполнение, что создает интерактивную учебную среду в группе.

4.4. Решение рабочих задач по теме автомобилизации.

Заключение

Интерактивное обучение и ИИ сегодня не просто инструменты – это интегрированная часть образовательной системы, которая делает обучение более персональным, позволят автоматизировать рутинные процессы, создавать качественный контент, методическое, информационное и программное обеспечение и использовать инновационные технологии.

Список использованных источников

1. Стратегия развития деятельности в области искусственного интеллекта в Республике Беларусь / А. М. Белоцерковский, В. В. Голенков, В. А. Головко [и др.]. – Минск : ОИПИ НАН Беларуси, 2024. – 8 с.

2. Капустин, В. В. Интерактивное обучение и практика : метод. пособие / В. В. Капустин. – Минск, 2024. – 26 с. – URL: https://autoexp.org/files/INTER-OBUCHENIE.pdf (дата обращения: 12.08.2025).

УДК 004.942, 519.876.5

Программный комплекс управления роботизированными системами на основе технологии цифровых двойников, методов генетической оптимизации и машинного обучения

Т. Ю. Ким
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: tkim@robotics.by

Введение

Современная робототехника переходит от жестко запрограммированных систем к автономным и обучаемым платформам. Важную роль в этом процессе играют цифровые двойники (ЦД) – виртуальные модели, позволяющие безопасно тестировать и оптимизировать поведение роботов без риска повреждения оборудования. ЦД находят применение в промышленности, медицине, транспорте и образовании.

Особенно актуальны цифровые двойники при работе с мобильными и человекоподобными роботами, отличающимися высокой нелинейностью и взаимодействием с динамичной средой. Прямое тестирование алгоритмов на физических прототипах требует значительных затрат и времени. Поэтому использование ЦД становится не просто удобным, а необходимым элементом современной инженерной практики.

Автором был разработан и реализован комплекс управления роботизированными системами на основе технологии цифровых двойников, методов генетической оптимизации (генетическими алгоритмами (ГА)) и машинного обучения (глубокое обучение с подкреплением, DRL, и подражательное обучением, IRL).

Основной целью исследований стало создание универсальной среды для проектирования, обучения и оптимизации поведения роботов с последующей адаптацией решений к реальным условиям.

Программный комплекс управления роботизированными системами

Разработка цифровых двойников роботизированных систем

Первым этапом работы стало построение точных имитационных моделей двух типов роботов:

– мобильный колесный робот с дифференциальными приводами;

– человекоподобный робот с 54 степенями свободы (верхние и нижние конечности).

Модели разрабатывались в среде MATLAB/Simulink с использованием Simscape Multibody, что позволило корректно учитывать физику движения, трение, демпфирование, инерцию и взаимодействие с датчиками.

Цифровой двойник использовался не только как инструмент визуализации, но и как полноценная испытательная площадка для алгоритмов управления. Это позволило проводить тысячи итераций обучения за короткое время, накапливать «опыт» в эквиваленте сотен часов реального времени и быстро заменять объект или метод с минимальной донастройкой [1].

Разработанная технология дает возможность создавать цифровые двойники, способные в режиме псевдореального времени имитировать поведение робота в различных условиях [2].

Оптимизация параметров с помощью генетических алгоритмов

На втором этапе был применен подход на основе ГА – эволюционного метода, имитирующего естественный отбор. ГА применялись для подбора коэффициентов ПИД-регулятора для мобильного робота и оптимизации траекторий движения по цветоконтрастной линии.

Были разработаны функции пригодности для различных типов трасс (окружность, эллипс, квадрат) и подобраны параметры операторов популяции, скрещивания и мутации. Каждая особь оценивалась через моделирование ее поведения в цифровом двойнике, что обеспечивало реалистичную обратную связь.

В результате анализа численных экспериментов установлены:

– рабочие значения коэффициентов для ПИД-регулятора мобильного робота передвигающегося по цветоконтрастной линии;

– минимальное влияние интегральной составляющей ПИД-регулятора;

– возможность упрощения системы управления до ПД-регулятора;

– оптимальные параметры популяции и числа итераций, минимизирующие вычислительные затраты.

Полученные коэффициенты успешно применены на учебном роботе RoboCake Beginner, который движется по цветоконтрастной линии со скоростью 0,4 м/с. Применение ГА с цифровым двойником позволило значительно сократить время настройки и повысить устойчивость управления [3].

Обучение с подкреплением для автономного управления

Третий этап был направлен на применение глубокого обучения с подкреплением для автономного принятия решений, позволяющего агенту самостоятельно формировать стратегию поведения с целью максимизации награды и минимизации штрафов.

Управление мобильным роботом

Для задачи следования по цветоконтрастной линии был реализован алгоритм TD3 (Twin Delayed DDPG, двойной отложенный глубокий детерминированный градиент политики).

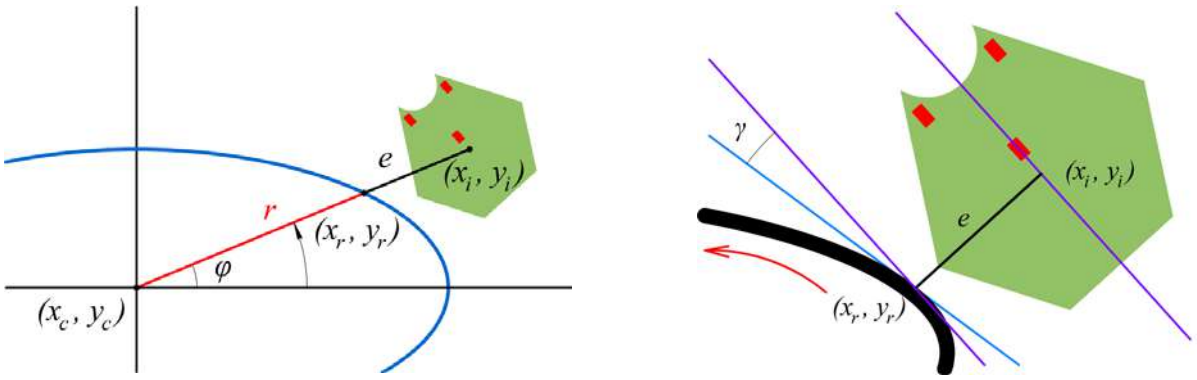
В качестве вознаграждения за выполненное действие учитываются два критерия. Во-первых, рассчитывается расстояние (e) как разность между расстоянием от центра робота (x_r, y_r) до центра эллипса и радиус-вектором эллипса (x_0, y_0) в данном направлении [6]:

$$e_i = \sqrt{(x_c - x_r^{(i)})^2 + (y_c - y_r^{(i)})^2} - r_i,$$

где x_c y_c – центр эллипса; $(x_r^{(i)}, y_r^{(i)})$ – координаты робота на i-м шаге; r_i – радиус-вектор эллипса в направлении φ_i, определяемом положением робота.

Чем ближе агент находится к краю эллиптической кривой, тем большее вознаграждение он получит (рис. 1, *а*).

Во-вторых, рассчитывается угол γ_i между направлением движения робота и касательной к эллипсу в точке $(x_r^{(i)}, y_r^{(i)})$ (рис. 1, *б*).



а) *б)*

Рис. 1. Графики, демонстрирующие формирование функции вознаграждения: *а*) вычисление угла φ расстояния, пройденного между осью абсцисс и текущей позицией робота с координатами x_r, y_r; r – радиус эллипса в текущий момент времени; *б*) определение минимального расстояния между кривой эллипса и центром робота (e) и вычисление угла γ между касательной к эллипсу и направляющей робота

Вектор скорости робота на шаге i определяется как

$$\vec{v}_i = (x_{i+1} - x_i, y_{i+1} - y_i).$$

Касательный вектор к эллипсу в точке $(x_r^{(i)}, y_r^{(i)})$ вычисляется по формуле

$$\vec{t_i} = \left(-\frac{y_r^{(i)}}{b^2}, \frac{x_r^{(i)}}{a^2} \right),$$

где a и b – полуоси эллипса.

Угол между векторами $\vec{t_i}$ и $\vec{v}$ определяется выражением

$$\gamma_i = \arccos\left(\frac{\vec{t_i} \cdot \vec{v_i}}{|\vec{t_i}| \cdot |\vec{v_i}|} \right).$$

Общая целевая функция, используемая в качестве награды, определяется как сумма этих двух компонентов:

$$R_{\text{line}} = e_i + \gamma_i,$$

где чем меньше значения e_i и γ_i, тем выше награда, получаемая агентом.

Реализован алгоритм TD3, который показал высокую эффективность в непрерывном пространстве действий. На основе алгоритма и функции вознаграждения построена модель, способная управлять роботом с минимальным отклонением от линии при скорости до 1 м/с.

Управление антропоморфным роботом

Для антропоморфного шагающего робота использовался алгоритм DDPG с функцией вознаграждения, учитывающей скорость, энергоэффективность и плавность движений.

Для обеспечения устойчивой имитации ходьбы была разработана функция вознаграждения, включающая положительную награду за линейную скорость движения вперед (v), что стимулирует продвижение робота. Для предотвращения отклонений от траектории и падений добавлены штрафы за отклонения по осям (y) и (z), а также за величину крутящего момента на звеньях (p) – это повышает энергоэффективность и сглаживает движения. Чтобы избежать раннего переобучения на быстром, но нестабильном поведении (например, падении вперед для увеличения скорости), введена дополнительная награда за продолжительность эпизода, стимулирующая сохранение равновесия и успешное выполнение задачи в течение длительного времени (t).

Итоговая функция вознаграждения имеет вид

$$r = w_1 v_y - w_2 \cdot p - w_3 z - w_4 x + w_5 t_s,$$

где $w_{1,\dots5}$ – весовые коэффициенты, определяющие относительную значимость каждой компоненты.

Общая суммарная награда R за эпизод рассчитывается как накопленная дисконтированная сумма (γ) мгновенных значений вознаграждения [4]:

$$R = \sum_{t=0}^{T} \gamma^t r_t.$$

Обучение проводилось на базе ЦД с использованием параллельных вычислений. После 6000 эпизодов робот демонстрировал стабильное передвижение вперед даже при наличии случайных возмущений [5]. Результаты подтверждают, что DRL позволяет формировать устойчивую походку без явного программирования траекторий, обеспечивая требуемую степень автономности и адаптивности.

Подражательное обучение с подкреплением

Четвертый этап – гибридный подход, сочетающий подражательное обучение и обучение с подкреплением. Он используется для быстрого старта обучения на основе экспертной демонстрации [6].

Был разработан метод, основанный на взаимодействии идеального мобильного робота (ИМР), следующего детерминированному алгоритму, и реального мобильного робота (РМР), обучаемого с помощью DRL.

Поставлена цель – обучить РМР следовать за ИМР, сохраняя центральное положение в коридоре, избегая столкновений и эффективно выходя из лабиринта. Функция вознаграждения учитывает следующие параметры: евклидово расстояние между роботами, данные лидара (расстояние до стен), линейную и угловую скорость, штрафы за столкновения и отклонение от курса, максимальную награду – за выход из лабиринта [7].

Для реализации использован алгоритм SAC (Soft Actor-Critic), который благодаря регуляризации энтропии обеспечивает более устойчивое и гибкое обучение. Эксперименты показали, что такой гибридный подход:

– существенно ускоряет обучение;

– повышает устойчивость в неизвестных средах;

– позволяет комбинировать детерминированные и стохастические стратегии.

Объединение в единую технологическую платформу

Все компоненты исследований объединены в единый технологический цикл, включающий построение цифрового двойника робота и его окружения, оптимизацию параметров с помощью ГА (настройку ПИД-регулятора), дообучение политик с использованием алгоритмов DDPG

и TD3, инициализацию поведения через IRL (SAC и демонстрации) и верификацию моделей в соответствии с выбранной системой управления.

Разработанная универсальная технология позволяет быстро создавать цифровые двойники мобильных и человекоподобных роботов, гибко выбирать методы управления в зависимости от задачи, адаптировать алгоритмы к изменяющимся условиям и интегрировать данные с датчиков и экспертных демонстраций. Это обеспечивает высокую эффективность разработки и тестирования автономных систем в виртуальной среде перед их внедрением на физические платформы [8].

Полученные результаты

Программный комплекс управления роботизированными системами на основе технологии цифровых двойников направлен на повышение эффективности разработки автономных роботов (рис. 2).

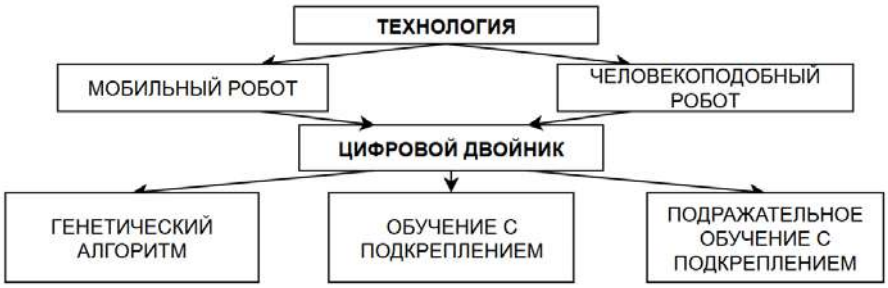


Рис. 2. Структура технологической платформы на основе цифровых двойников для управления роботизированными системами

Основные результаты:

– создана универсальная технология создания цифровых двойников для мобильных и человекоподобных роботов;

– реализовано применение генетических алгоритмов, DRL (DDPG, TD3) и подражательного обучения (SAC) в единую систему;

– сокращено время разработки и повышена устойчивость алгоритмов управления.

Создан программный комплекс, применимый в научных, образовательных задачах. Использование данного комплекса позволяет снизить временные и финансовые затраты на тестирование, повысить надежность решений и обеспечить высокую адаптивность систем.

Заключение

Разработанный программный комплекс представляет собой комплексное решение для повышения эффективности создания автономных роботов. Применение цифровых двойников в сочетании с методами генетической

оптимизации и машинного обучения позволяет создавать универсальные системы управления, устойчивые к изменениям внешней среды. Цифровые двойники обеспечивают быструю настройку, тестирование и верификацию алгоритмов в условиях моделирования, что снижает зависимость от физических прототипов. Полученные решения внедрены в научную практику и подтверждены актами о внедрении.

Технология применима в образовательных, исследовательских и промышленных задачах, связанных с управлением сложными динамическими системами.

Благодарность. Работа выполнена при поддержке гранта БРФФИ Ф23МЭ–030 «Эволюционные методы генерации оптимальных структур высокопроизводительных аппаратных ускорителей для реализации искусственных нейронных сетей на базе перепрограммируемых логических интегральных схем» в рамках задания 1.3.1. ГПНИ «Цифровые и космические технологии, безопасность человека, общества и государства», подпрограмма «Цифровые технологии и космическая информатика».

Список использованных источников

1. Ким, Т. Ю. Разработка цифрового двойника мобильного робота для исследовательских и учебных целей на базе MATLAB/SIMULINK / Т. Ю. Ким, Г. А. Прокопович // Молодежь в науке – 2021: тез. докл. XVIII Междунар. науч. конф. молодых ученых, Минск, 27–30 сент. 2021 г. : в 2 ч. Ч. 2. Медицинские, физико-математические, физико-технические науки, химия и науки о Земле / Нац. акад. наук Беларуси, Совет молодых ученых; редкол.: В. Г. Гусаков (гл. ред.) [и др.]. – Минск : Беларус. навука, 2021. – С. 580–584.

2. Ким, Т. Ю. Разработка цифрового двойника антропоморфного робота «астронавт» с применением метода обучения с подкреплением в режиме реального времени / Т. Ю. Ким // Молодежь в науке / НАН Беларуси. – Минск, 2022. – С. 580–584.

3. Ким, Т. Ю. Оптимизация коэффициентов ПИД-регулятора системы управления движением мобильного робота по цветоконтрастной линии на основе генетического алгоритма / Т. Ю. Ким, Г. А. Прокопович // Информатика. – 2021. – Т. 18, № 4. – С. 53–68. – DOI: 10.37661/1816–0301–2021–18–4–53–68.

4. Kim, T. Automatic Tuning of the Motion Control System of a Mobile Robot Along a Trajectory Based on the Reinforcement Learning Method / T. Kim, R. Prakapovich // Communications in Computer and Information Science. – Springer, Cham, 2022. – Vol. 1562. – P. 234–244. – DOI: 10.1007/978–3–030–98883–8_17.

5. Kim, T. Y. Control of an Anthropomorphic Walking Robot Using Reinforcement Learning // Embedding the Future : 17th Intern. Conf. on Pattern Recognition and Information Processing, Minsk, Belarus. – Minsk : UIIP NASB, 2025. – P. 147–151.

6. Ким, Т. Ю. Обучение мобильного робота передвижению в неизвестной среде с применением метода обучения с подкреплением / Т. Ю. Ким // Информационные технологии в промышленности, логистике и социальной сфере» (ITI*2023) : тез. докл. 12-й Междунар. науч.-техн. конф. / Объед. ин-т проблем информатики Нац. акад. наук Беларуси ; под науч. ред. А. В. Тузиков, П. Н. Бибило, А. Г. Гривачевский. – Минск : ОИПИ НАН Беларуси, 2023. – С. 78–81.

7. Ким, Т. Ю. Разработка метода подражательного обучения для нейросетевой системы управления движением мобильного робота на примере задачи поиска выхода из лабиринта / Т. Ю. Ким, Г. А. Прокопович // Информатика. – 2024. – Т. 21, № 3. – С. 32–46. – DOI: 10.37661/1816–0301–2024–21–3–32–46.

8. Ким, Т. Ю. Технология создания цифровых двойников для оптимизации конструктивных параметров роботов и их систем управления / Т. Ю. Ким, Г. А. Прокопович // Цифровая трансформация. – 2025. – Т. 31, № 3. – С. 16–23.

УДК 37.022; 004.75

Трансформация цифровой грамотности в ИИ-компетентность как ключевой ресурс адаптации преподавателя в цифровой образовательной среде

Н. Л. Ковалёва
Республиканский институт профессионального образования, Минск, Беларусь
E-mail: kolkova7@mail.ru

А. П. Ковалёв
Белорусский государственный экономический университет, Минск
E-mail: alex_kov2001@mail.ru

Введение

В настоящее время цифровая трансформация вносит революционные изменения в образовательные процессы, идет такими быстрыми темпами, что учреждения образования и преподаватели не всегда успевают перестраиваться и оперативно реагировать на перемены [1, с. 137]. Поэтому важной задачей системы образования является обеспечение кадрового потенциала, необходимого для этой трансформации.

Цифровая трансформация системы образования Республики Беларусь закреплена в ряде стратегических документов, таких как Концепция обеспечения суверенитета в сфере цифрового развития и Программа социально-экономического развития [2, 3], и выступает системообразующим фактором, требующим пересмотра подходов к подготовке педагогических кадров.

Динамичное развитие и внедрение технологий искусственного интеллекта (ИИ) вносят свои коррективы в понимание цифровых компетенций преподавателя, одновременно создавая вызовы и открывая новые перспективы. Так, декан экономического факультета МГУ А. Аузан сказал об ИИ в образовании: «ИИ распространяется как “торфяной пожар” – скрыто, но неизбежно. Можно предположить, что нынешняя система образования из-за этого изменится даже быстрее, чем рынок труда. В ближайшие 3–5 лет она просто “сгорит” с привычными для нас контрольными, тестами и экзаменами».

Очевидно, что, обладая потенциалом изменить способ обучения учащихся и преподавателей, ИИ преобразует будущее образования, а пока

даже не вписывается в традиционные образовательные категории. Это не инструмент, не учитель и не учебник (электронный учебно-методический комплекс), а новый участник (субъект образовательного процесса), который пока не описан в педагогической теории [4].

Все вышеизложенное говорит о необходимости трансформации образования, перехода из области теоретизирования в область практического применения полученных знаний. Хотя такой переход упирается в эволюцию самих образовательных моделей.

Традиционная модель информационной культуры, включающая работу с данными, критическую оценку источников и создание контента, сегодня становится необходимой, но уже недостаточной основой. Именно на этом фоне явно проявляется существующий разрыв. Наблюдается кризис профессиональной компетентности, выражающийся в неготовности многих преподавателей адаптироваться к новым технологиям, а также их фрагментарном использовании в образовательном процессе.

Кризис традиционной модели

История последних двадцати лет показывает, что ожидания технологических революций в образовании часто не оправдываются. Каждый технологический цикл – от системы управления обучением (Learning Management System, LMS) и массовых открытых онлайн-курсов (Massive Open Online Courses, MOOCs) до массового перехода на облачную платформу для проведения видеозвонков, конференций, вебинаров и онлайн-занятий (Zoom) во время пандемии – проходил по одинаковой схеме.

2000-е гг.: e-learning 1.0 (электронное обучение) и LMS. Было обещано, что корпоративные LMS (платформы Moodle, Blackboard и др.) изменят образование, все процессы перейдут в онлайн-пространство. Но LMS стали стандартом, преподавание не изменилось, материалы стали доступны в режиме онлайн.

2010-е гг.: MOOC-революция (платформы Coursera, edX и др.). Предполагалось, что придет бесплатное онлайн-образование для всего мира, произойдет замена университетов, миллионы студентов из Африки и Индии будут учиться у профессоров Стэнфорда. Произошло то, что MOOCs стали популярны среди специалистов ИТ-сферы и самообучающихся, но реальная вовлеченность пользователей и заинтересованных лиц составила порядка 10 %. Университеты как были, так и остались. MOOCs встроились в рынок как дополнение в виде корпоративных подписок, сертификации и др.

2015 г.: мобильное обучение и микро-learning. Предполагалось, что мобильные приложения заменят LMS, обучение микрообучением, доступным везде. Но мобильные курсы стали нормой, они не вытеснили классические форматы. Большинство LMS внедрили мобильные клиенты.

2017–2020 гг.: VR/AR (виртуальная и дополненная реальность) и «метавселенные» для обучения. Предполагалось, что очки виртуальной реальности заменят аудитории, обучение станет иммерсивным. Оказалось, что технологии являются дорогими и громоздкими. В итоге VR используют для нишевых тренингов, например в медицине или пожарной безопасности, но массового применения не наблюдается.

2020 г. (COVID): Zoom-революция. Предполагалось, что раз уж образовательный процесс ушел в Zoom (онлайн-формат), то образование навсегда станет гибридным и цифровым. Онлайн-формат закрепился, но качество в основном упало, лекции стали просто «говорящей головой в Zoom». После пандемии большинство заинтересованных лиц и пользователей вернулось к офлайн или гибриду.

2023 г.: ИИ-революция (ChatGPT, RAG (Retrieval Augmented Generation – генерация с дополненной выборкой), ассистенты). Предполагается, что ИИ заменит преподавателей, индивидуализирует обучение, создаст новые курсы за минуты, перестроит HR-обучение. Реально происходит то, что быстрее создаются тесты, задания, переводы, появляются чатботы-ассистенты, но LMS, курсы и подходы остаются все теми. Преподаватели и HR учатся пользоваться промптами (запросами в контексте), не меняя саму педагогику.

Из данного хронологического обзора за последние двадцать лет можно сделать вывод, что каждый цикл идет по одной схеме: повышенное внимание к теме → ожидание революции → массовое разочарование → тихая интеграция как дополнение. То, что работает – остается (LMS, MOOC, мобильные приложения, Zoom, ИИ-ассистенты), но существенной революции не случилось. Этот опыт доказывает, что сами по себе технологии не производят переворот. Ключевой фигурой трансформации была и остается фигура преподавателя. Именно от его компетенций зависит, станет ли цифровая трансформация прогрессивным шагом или обернется серьезными рисками.

Цифровая трансформация – это не просто добавление цифровых инструментов, а процесс, в результате которого традиционные способы организации деятельности полностью заменяются новейшими цифровыми альтернативами. Однако такой процесс сопряжен с рисками: системными проблемами в сфере образования, размытием границ личного и рабочего времени и даже угрозой депрофессионализации преподавательского труда. Таким образом, готовность преподавателя к работе в новых условиях становится центральной проблемой, требующей детального анализа.

Разрыв между требованиями цифровой эпохи и реальными возможностями педагогов обусловливает целесообразность перехода от понятия цифровой грамотности к более сложной модели. Появляется необходимость в создании концептуальной рамки для такой модели с выделением

в ней пороговых и дифференцирующих компетенций, а также определении места новой, так называемой ИИ-компетентности как ключевого фактора профессионального превосходства современного преподавателя. Авторы, используя термин «цифровая грамотность», вкладывают в него единый смысл, который подразумевает возможности эффективного использования цифровых технологий в преподавании и обучении [5].

Новая парадигма профессионализма

В условиях технологической реальности целесообразно переосмысливать компетентностную модель через призму ее разделения на два уровня: пороговые компетенции как базовый уровень и дифференцирующие компетенции как фактор профессиональной исключительности.

В пороговые компетенции включается минимально необходимый набор знаний, умений и навыков, без которых невозможна эффективная педагогическая деятельность в современной цифровой среде. Они представляют собой фундамент, на котором строятся более сложные профессиональные навыки. К ним относятся: информационно-коммуникационная грамотность, цифровая дидактика, кибербезопасность и цифровая этика.

Эти компетенции являются обязательным минимумом, который обеспечивает базовую готовность преподавателя к работе в цифровом образовательном пространстве.

В дифференцирующие компетенции включаются навыки более продвинутого уровня, выделяющие лучших специалистов и определяющие их конкурентоспособность и уникальность. Именно они позволяют преподавателю не просто использовать технологии, а трансформировать образовательный процесс. В контексте эпохи ИИ можно выделить следующие ключевые дифференцирующие компетенции: надпрофессиональные и метакомпетенции, а также ИИ-компетентность.

К надпрофессиональным и метакомпетенциям относятся критическое мышление в цифровой среде (способность анализировать ИИ-контент, цифровые данные, выявляя дезинформацию и манипуляции), адаптивность и самообучение, креативность, а также предусмотрительность (способность предвидеть технологические тренды и адаптировать образовательный процесс к будущим изменениям, работая на опережение).

ИИ-компетентность – это компетенция, включающая способность эффективно и этично применять ИИ-инструменты для решения педагогических задач. Она предполагает не просто использование готовых сервисов, а глубокое понимание их возможностей и ограничений. В структуру ИИ-компетентности входят: педагогический дизайн с ИИ, эффективный промпт-инжиниринг, критическая оценка ИИ-контента и аналитика образовательных данных.

Такая ИИ-компетентность как педагогический дизайн с ИИ предполагает, что преподаватель должен уметь проектировать учебные сценарии и задания, которые дополняются, а не заменяются ИИ. Педагогический дизайн включает в себя создание симуляций, адаптивных тестов и персонализированных образовательных траекторий.

Эффективный промпт-инжиниринг подразумевает, что преподаватель должен иметь навык правильной формулировки запросов в контексте к нейросетям для получения качественных образовательных материалов (планов занятий, кейсов, визуального контента и т. д.).

Критическая оценка ИИ-контента предполагает умение преподавателем выявлять и корректировать ошибки (галлюцинации), предвзятость и неточности в материалах, сгенерированных ИИ, а также обучать этому учащихся.

Аналитика образовательных данных в ИИ-компетентности определяет умение преподавателя интерпретировать собранные ИИ-системами данные об успеваемости для своевременной коррекции образовательного процесса и выявления индивидуальных трудностей учащихся.

Освоение дифференцирующих компетенций ведет к трансформации профессиональной роли. Преподаватель перестает быть просто транслятором знаний и становится цифровым наставником, изменяются его функции. Ключевые функции смещаются в сторону проектирования персонализированных образовательных траекторий. Используя аналитику ИИ, преподаватель помогает каждому учащемуся выстраивать свой уникальный путь обучения. Также появляется функция курирования цифровых ресурсов, когда преподаватель выступает как эксперт, который отбирает, адаптирует и рекомендует качественные цифровые и ИИ-инструменты.

Следующая функция формирования цифровой культуры и ИИ-грамотности включает в себя ответственное, этичное и критическое использование технологий при обучении учащихся, формируя навыки, необходимые для жизни и работы в XXI в.

Функция фокусировки на человеческом взаимодействии определяет, что, освободившись от рутинных задач благодаря ИИ, преподаватель может уделять больше времени мотивации, развитию «мягких» навыков, воспитательной работе и живому общению, это невозможно заменить технологиями. Есть и обратная сторона такого взаимодействия, заключается она в ряде педагогических рисков.

Риски и вызовы интеграции ИИ

Несмотря на огромный потенциал, внедрение ИИ в образовательный процесс порождает педагогические риски и вызовы. Основной риск заключается в снижении когнитивных способностей. Главная опасность – ухуд-

шение критического мышления и формирование поверхностных знаний. По исследованию, проведенном среди студентов-программистов, было отмечено, что при успешном выполнении практических работ с помощью ИИ, уровень теоретических знаний остается низким.

Опыт некоторых организаций показывает, что если «кормить» нейросеть только существующей внутренней информацией, то система в какой-то момент может замкнуться сама на себе, ее развитие останавливается, а сотрудники попадают в ловушку ИИ. Чтобы избежать этого, важно сохранять независимый от ИИ интеллектуальный центр. В образовании таким центром является преподаватель, который будет «подкармливать» алгоритмы новыми данными.

Внедрение ИИ в образовательный процесс порождает этические и технические вызовы. Ключевыми проблемами являются: конфиденциальность данных учащихся, цифровое неравенство; вопросы академической честности и авторского права; риск алгоритмической предвзятости, проявляющийся в систематической ошибке в работе алгоритмов ИИ, что приводит к несправедливым, дискриминационным или необъективным результатам в отношении отдельных групп людей.

Заключение

Интеграция ИИ в образование – это не просто внедрение новых инструментов, а сложный процесс трансформации, требующий пересмотра модели профессиональной компетентности преподавателя. Разделение на пороговые и дифференцирующие компетенции позволяет создать более гибкую и адаптивную систему профессионального развития, отвечающую вызовам времени.

Формирование ИИ-компетентности и связанных с ней метанавыков становится важным фактором конкурентоспособности как отдельного преподавателя, так и системы профессионального образования в целом. Конечная цель – не заменить человека машиной, а построить эффективную команду в связке «человек и ИИ», где технологии усилят педагогическое мастерство, а преподаватель выступит в роли наставника в сложном цифровом мире.

Перспективным направлением дальнейших исследований являются разработка диагностического инструментария для оценки уровня сформированности дифференцирующих компетенций и создание на его основе моделей непрерывного повышения квалификации, ориентированных на подготовку цифровых наставников будущего. Исследования в этой области позволят более полно раскрыть потенциал ИИ и разработать эффективные стратегии его внедрения в учреждения образования, обеспечивая баланс между инновациями и традиционными методами обучения.

Список использованных источников

1. Ковалёв, А. П. Повышение эффективности образовательного процесса в учреждениях высшего образования с использованием искусственного интеллекта / А. П. Ковалёв, Н. Л. Ковалёва // Искусственный интеллект в Беларуси : доклады III Форума IT-Академграда, Минск, 10-11 окт. 2024 г. – Минск : ОИПИ НАН Беларуси, 2024. – С. 137–142.

2. О Концепции обеспечения суверенитета Республики Беларусь в сфере цифрового развития до 2030 года (в ред. постановления Совета Министров Республики Беларусь от 31 декабря 2024 г. № 1074 / Зарегистрировано в Национальном реестре правовых актов Республики Беларусь, 15.01.2025, 5/54484).

3. Программа социально-экономического развития Республики Беларусь на 2021–2025 годы (в ред. постановления Совета Министров Республики Беларусь от 23 июня 2023 г. № 180 / Зарегистрировано в Национальном реестре правовых актов Республики Беларусь, 28.06.2023, 1/20899).

4. Kovaleva, N. The role of artificial intelligence in the formation of the educational process future / N. Kovaleva // Развитие образования посредством искусственного интеллекта: материалы Междунар. науч.-практ. конф., 20 сент. 2024 г., Казань. – Казань : Университет управления «ТИСБИ», 2024. – С. 3–8.

5. Ковалёва, Н. Л. Формирование цифровых компетенций преподавателя в условиях трансформации образовательной среды / Н. Л. Ковалёва // Актуальные вопросы профессионального образования = Actual issues of professional education : материалы VI Междунар. науч.-практ. конф., г. Минск, 12–13 июня 2025 г. / УО «Белорусский государственный университет информатики и радиоэлектроники», 2025. – С. 224–228.

УДК 004.8.032.26; 004.93'1; 004.932; 004.89:004.4; 616-006

Обнаружение узловых образований на КТ-изображениях легких на основе методов глубокого обучения

А. А. Косарева, Э. В. Снежко✉, В. А. Ковалев
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: eduard.snezhko@gmail.com

Введение

Согласно статистическим данным, опубликованным Национальным институтом рака США [1], рак легких является третьим по частоте диагностирования видом рака и ведущей причиной смерти от рака в 2024 г. Раннее диагностирование рака легких у пациентов приобретает все большее значение. Согласно статистическим данным у лиц с диагнозом рака легких на ранней стадии вероятность выживания в течение как минимум пяти лет после постановки диагноза значительно выше, чем у лиц с диагнозом на более поздней стадии.

Было доказано, что использование компьютерной томографии в качестве диагностического инструмента способствует снижению смертности от рака легких [2, 3]. Разработка алгоритмов компьютерной диагностики (CAD) для выявления рака легких на ранней стадии может повысить эффективность работы радиологов в диагностике этого заболевания на начальных стадиях. Узловые образования имеют средний диаметр до 3 мм и достигают диаметра примерно 56 мм при развитии. Относительно небольшой размер рассматриваемых узловых образований усложняет процесс диагностики рака. Следовательно, при разработке алгоритмов их автоматического обнаружения важно учитывать этот фактор.

Проблема раннего выявления узловых образований в легких уже тщательно изучена исследователями из многочисленных научно-исследовательских центров. Использование алгоритмов глубокого обучения обеспечивает прочную основу для соответствующих предлагаемых методов и алгоритмов. Например, в работе [4] авторы использовали сверточные нейронные сети для выявления метастазов в легких на изображениях микрокомпьютерной томографии (микро-КТ) мышей. В исследовании показатели площади под кривой (AUC) достигли значений (0,76–0,77) на обычном наборе данных и улучшились на 8 % при разбавлении набора данных искусственно сгенерированными данными.

В своей работе авторы [5] использовали нейронную сеть Faster R-CNN для выявления потенциальных кандидатов на снимках компьютерной томографии легких. Выявленные области были впоследствии классифицированы с помощью трехмерной сверточной нейронной сети, разработанной на основе многомасштабных блоков внимания. В работе [6] в качестве детектора области интереса использовалась модель YOLOv5s с пороговым значением вероятности обнаружения объекта 0,3. Низкий порог вероятности был использован для обеспечения обнаружения большего количества подозрительных объектов. Для определения типа идентифицированных объектов (т. е. наличия или отсутствия узловых образований) был использован трехмерный классификатор на основе сверточной нейронной сети. Использование технологии YOLO привело к значительному сокращению времени, необходимого для анализа одного снимка КТ легких.

Все рассмотренные выше подходы были ориентированы на выявление узловых образований в легких и не затрагивали вопрос определения доброкачественного или злокачественного характера выявленных образований. С другой стороны, в исследовании [7] авторы рассмотрели вопрос определения злокачественности обнаруженных узлов. Для этого они использовали нейронную сеть, переработанную для трехмерных изображений, и метод HESAM [8] с мягким отображением активации, что позволило проводить анализ изображений на детальном уровне. Совмещение обнаруженных областей интереса с векторами признаков фрагментов КТ-изображений, извлеченных посредством интерпретации карт активации, показало себя эффективным подходом для классификации узловых образований в соответствии с их злокачественностью. Карты активации были получены как выходные значения на определенных слоях нейронной сети во время выполнения ею задачи сегментации узловых образований.

Данные и методы

Данные

Обучение нейронных сетей проводилось с использованием трех наборов данных для двух задач: локализации узлов и их разделения на доброкачественные и злокачественные.

1. Набор данных LUNA16 содержал 1186 узлов в легких на 888 компьютерных томограммах легких, которые были аннотированы четырьмя радиологами (ds1). Аннотация включала в себя расположение областей и метку «узел/не узел».

2. Внутренний набор данных содержал 323 аннотированных узла на 200 компьютерных томограммах легких (ds2). Аннотация включала в себя расположение узлов и метку «злокачественный» или «доброкачественный».

Предлагаемый алгоритм обнаружения и классификации узлов состоял из трех этапов:

1. Обнаружение областей интереса (ROI) на 2D-аксиальных срезах 3D-КТ-изображений легких. Области интереса, обнаруженные на этом этапе, содержали потенциальные узлы.

2. Бинарная классификация областей интереса: «узел» или «не узел».

3. Бинарная классификация обнаруженных узлов с целью дифференциации злокачественных и доброкачественных образований.

На этапе обнаружения ROI для выявления потенциальных узлов использовались двумерные срезы. Размер срезов составляли 512×512 пикселей. Динамический диапазон значений пикселей был ограничен значениями [-1500, 500] для лучшей визуализации структур в легких [9]. Для задач бинарной классификации использовались трехмерные изображения с динамическим диапазоном значений от [-1500, 500], нормализованные в диапазоне от [-1, 1] и масштабированные таким образом, чтобы все коэффициенты матрицы преобразования были равны друг другу.

Методы

В качестве архитектуры для обнаружения областей интереса на КТ-изображениях легких была выбрана нейронная сеть YOLOv8 [10]. Это позволяет обнаруживать все потенциальные узлы за один проход во время сетевого вывода.

На этапе тренировки нейронной сети на обучающем наборе данных отслеживались следующие показатели: пересечение над объединением (IoU), средняя точность (AP) и средняя точность (mAP) по различным классам объектов. Таким образом, mAP50 рассчитывался при пороговом значении IoU = 0,50, mAP50-95 – при пороговом значении IoU от 0,50 до 0,95. Было предложено обучить нейронную сеть YOLOv8 на наборе данных, разделенном на два поднабора в зависимости от размера узлов.

После локализации областей-кандидатов реализована бинарная классификация «узел», «не узел» для фильтрации ложных срабатываний.

На последнем этапе была использована нейронная сеть из семейства DenseNet для идентификации и удаления областей без узлов. В этом конкретном сценарии трехмерные фрагменты КТ-изображений извлекались с использованием координат ранее идентифицированных кандидатов.

Результаты

Обнаружение узловых образований

Для задачи обнаружения узлов была выбрана модель YOLOv8s. Было проведено исследование эффективности предварительно обученных моделей с разным количеством параметров, которое показало, что модели с большей сложностью не улучшали качество обнаружения.

При обучении нейронной сети были выбраны следующие параметры:

1. Классификация узлов была отделена от этапа обнаружения. В целях минимизации влияния классификации на чувствительность модели детектирование выполнено отдельно на отобранных кандидатах.

2. Параметры mosaic и mixup модели YOLOv8s были установлены на ноль. Такой выбор был обусловлен спецификой задачи. Увеличение изображений путем комбинирования различных частей было сочтено ненужным, так как это привело бы к искажению анатомических особенностей.

3. Образцы были случайным образом повернуты в диапазоне до 30 град.

4. В некоторых случаях образцы были случайным образом отражены по двум осям.

5. Образцы были масштабированы с коэффициентом от 0,0 до 0,3.

При настройке архитектуры нейронной сети для задачи обнаружения узлов были учтены два ключевых фактора. Во-первых, набор данных был разбавлен срезами КТ, на которых отсутствовали обнаруживаемые объекты (так называемые фоновые слои). Во-вторых, набор данных был дополнен КТ-срезами, соседствующими с теми, на которых присутствует аннотация. Для обеспечения целостности набора данных все срезы были вручную проверены с целью выявления и исключения любых ошибочных данных. Внедрение этих усовершенствований привело к значительному повышению эффективности обнаружения узлов.

В ходе экспериментов, проведенных на обучающих и тестовых наборах данных, было отмечено, что на результат обнаружения в тестовом наборе данных значительно влияют размер обнаруженных узлов и состав обучающих и тестовых образцов. Таким образом, изображение одного пациента может быть включено в обучающий, валидационный или тестовый наборы. Размер узлов в наборах оказал значительное влияние на результат. Было отмечено, что нейронная сеть продемонстрировала недостаточную способность точно обнаруживать как крупные, так и мелкие образования. Этот результат стал основой для создания нового алгоритма поиска, связанного с размерами детектируемых узлов. Наборы данных были разделены в соответствии с нормализованными размерами узлов. Пороговое значение, используемое для классификации образцов изображений, содержащих узлы, как «большие» или «малые», было определено как 0,04. Это значение получено не только на основе среднего значения по набору данных, но и на основе количества элементов в каждом поднаборе данных в процессе разбиения.

Двойное обнаружение объектов продемонстрировало заметную эффективность, дав весьма обнадеживающие результаты. Использование двух отдельно обученных моделей дало значение точности 0,941 по тестовому набору данных, полноты 0,900 и f1-меры 0,920. Тем не менее сети проде-

монстрировали способность идентифицировать большее количество объектов, чем ожидалось. Поэтому было решено, что внедрение вспомогательного классификатора послужит повышению точности классификации объектов.

Еще одной ключевой задачей было эмпирическое определение порога достоверности обнаружения. От этого значения зависело количество найденных узлов, а также количество ошибочно обнаруженных объектов. Порог достоверности обнаружения определялся путем оценки доли «ненайденных» узлов и количества ложноположительных обнаружений. Оптимальный порог достоверности для обнаружения «мелких» узлов был определен как 0,2, а для «крупных» узлов – как 0,4.

Классификатор «узел/не узел»

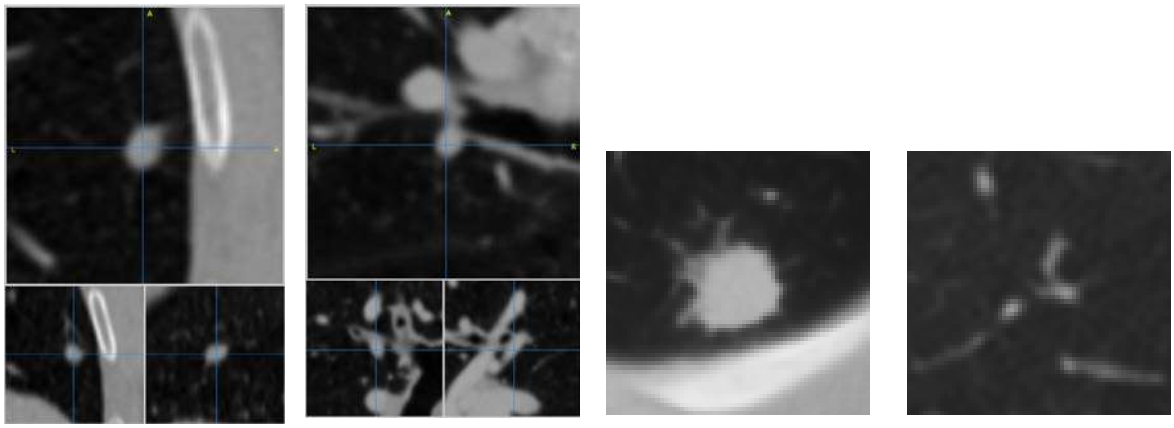
Классификатор узлов был разработан для отделения ложноположительных объектов, некорректно обнаруженных на 2D-срезах КТ. Для каждого кандидата была извлечена 3D-область, окружающая соответствующий фрагмент, что значительно расширило объем доступной информации и упростило определение, представляет ли собой кандидат истинное узловое образование. Основной гипотезой являлось предположение, что внедрение классификатора трехмерных фрагментов КТ-изображений повысит точность обнаружения за счет более надежного отсеивания артефактов.

К ложноположительным объектам относились сосуды, фрагменты других органов и тканей, не связанные с патологическими новообразованиями, а также прочие включения. В рамках набора данных ds1 объекты, помеченные как «кандидаты» с «классом», равным 0, использовались в качестве негативных примеров (не узлы) при обучении модели. Их координаты были автоматически извлечены с помощью алгоритма, основанного на пороговой фильтрации.

Было принято решение обучить нейронные сети DenseNet для решения задачи классификации трехмерных образцов и YOLOv8 для решения задачи классификации двухмерных образцов. Такой подход был выбран для подтверждения гипотезы о преимуществах работы с трехмерными объектами.

На следующих иллюстрациях представлены примеры изображений, входящих в обучающие наборы данных (рис. 1). Трехмерные фрагменты представляли собой кубы с равным расстоянием между осями.

Результаты классификации на тестовом наборе данных представлены в табл. 1. В данном исследовании использовались модели DenseNet121 и YOLOv8n, которые продемонстрировали способность обеспечивать высококачественную классификацию без переобучения.



а) *б)* *в)* *г)*

Рис. 1. Примеры изображений в каждом из наборов данных: *а*) 3D-узел; *б*) 3D-не узел: *в*) 2D-узел, *г*) 2D-не узел

Таблица 1

Результаты классификации на классы «узел» и «не узел»

	DenseNet121			YOLOv8n		
Метрика	Точность	Полнота	f1-мера	Точность	Полнота	f1-мера
Значение	0,971	0,971	0,971	0,971	0,900	0,934

Как и ожидалось, модель DenseNet121, обученная на трехмерных фрагментах КТ-изображений, показала лучшие результаты в задаче классификации «узел»/«не узел». В то же время YOLOv8n продемонстрировала склонность к ложноотрицательным ошибкам, пропуская ряд потенциально значимых объектов. Тем не менее обе нейронные сети подтвердили свою эффективность и применимость для решения поставленной задачи.

Разработка классификатора злокачественных или доброкачественных новообразований

Как и в случае с предыдущим этапом, было принято решение сравнить результаты обучения классификатора трехмерных образцов и классификатора двухмерных образцов. Однако в данном случае задача оказалась более сложной. Были выявлены следующие проблемы:

1. Небольшой объем данных для классификатора трехмерных образцов.

2. Данные были получены из набора данных ds2. В результате проверка результата оказалась сложной задачей.

3. Глубина модели DenseNet121 оказалась недостаточной для целей данного исследования, что потребовало включения DenseNet169 и DenseNet264, которые имеют большее количество обучаемых параметров.

Для повышения точности классификации двухмерных образцов были сформированы трехслойные трехмерные изображения. Каждый соседний слой был включен как один из трех каналов изображения (рис. 2).

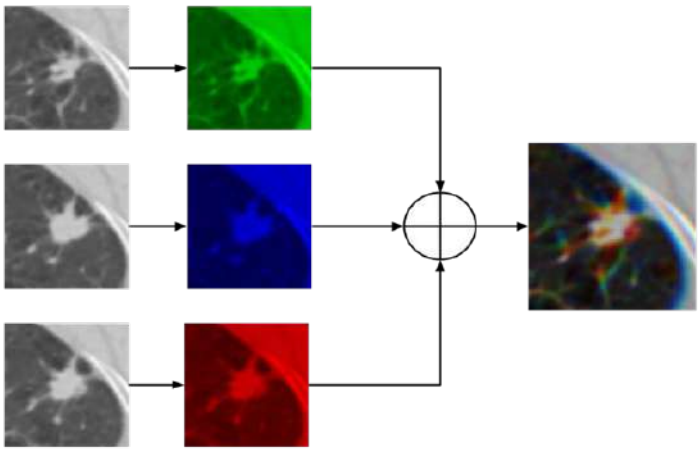


Рис. 2. Метод формирования псевдотрехмерных изображений из трех срезов

Результаты экспериментов представлены в табл. 2. Набор данных из двухмерных образцов продемонстрировал более высокие значения показателей благодаря возможности извлечения нескольких слоев из одного трехмерного узла. Изображения узлов отдельных пациентов были разделены на три группы: тестовые, валидационные и обучающие.

Таблица 2

Результаты классификации «злокачественные» или «доброкачественные»

	DenseNet264 (3D)			YOLOv8n (2D)			YOLOv8n (три слоя)		
Метрика	Точность	Полнота	f1-мера	Точность	Полнота	f1-мера	Точность	Полнота	f1-мера
Значение	0,787	0,791	0,789	0,655	0,731	0,691	0,840	0,952	0,892

Как показано в таблицах, классификация трехмерных образцов с использованием трех соседних срезов КТ-изображений продемонстрировала лучший результат. Тестовый набор данных не содержал значительного количества данных для классификации трехмерных образцов. Хотя выводы, сделанные нейронной сетью, могут показаться произвольными, важно отметить, что нейронная сеть DenseNet264 была обучена три раза и отклонение в показателях каждый раз находилось в пределах 2–3 %. Классифика-

ция двухмерных образцов показала худшие результаты по сравнению с классификацией трехмерных образцов.

В ходе исследования основным недостатком при выполнении классификации трехмерных образцов оказалось ограничение, связанное с объемом доступного набора данных. В то же время ключевым преимуществом классификации двухмерных образцов стала возможность расширения обучающей выборки за счет использования соседних слоев КТ-изображений. Основным достоинством трехмерного подхода являлось более полное извлечение информации о пространственной структуре узла, что способствовало лучшему пониманию его объемной морфологии.

Использование трехслойных образцов позволило сбалансировать два метода и дало лучшие результаты.

Заметным преимуществом использования трехслойных фрагментов КТ-изображений являлась возможность точного обнаружения узлов малого диаметра в пределах трех-пяти слоев. Сложный характер трехмерного изображения, состоящего из 64 слоев, мог привести к наличию избыточной информации.

В целом задача классификации узлов на злокачественные и доброкачественные оказалась более сложной по сравнению с классификацией «узел/не узел». Даже для медицинского специалиста может быть сложно определить, обладает ли конкретный узел злокачественными или доброкачественными характеристиками.

Заключение

Целью исследования являлась разработка алгоритмов и обучение соответствующих нейронных сетей, предназначенных для поддержки раннего выявления узловых образований на компьютерных томограммах легких. Данный подход направлен на выявление подозрительных областей, требующих дополнительного анализа, а также на оценку вероятности злокачественности обнаруженных образований. Предполагается, что это позволит сократить время анализа изображений специалистом и облегчит процесс интерпретации КТ-данных.

Качество обнаружения объектов было улучшено за счет разделения обучающего набора данных на поднаборы, соответствующие определенным размерам узлов. Классификатор трехслойных образцов, сформированных из трех соседних срезов КТ-изображений, на основе сети YOLOv8n продемонстрировал лучшую эффективность в оценке злокачественности обнаруженных узлов, в то время как классификатор трехмерных образцов продемонстрировал повышенную эффективность в различении узлов от прочих анатомических структур.

Список использованных источников

1. SEER Stat Fact Sheets: Lung and Bronchus Cancer // National Cancer Institute. – URL: http://seer.cancer.gov/statfacts/html/lungb.html (date of access: 05.02.2024).

2. Reduced Cancer Mortality with Low-Dose Computed Tomographic Screening / The National Lung Screening Trial Research TeamAuthor Info & Affiliations. – 2011. – URL: http://www.nejm.org/doi/full/10.1056/NEJMoa1102873#t=article (date of access: 05.02.2024). – DOI: 10.1056/NEJMoa1102873.

3. Reduced Lung-Cancer Mortality with Volume CT Screening in a Randomized Trial / J. H. de Koning, C. M. van der Aalst, P. A. de Jong [et al.] // New England Journal of Medicine. – 2020. – Vol. 382, no. 6. – P. 503–515.

4. Detection of Lung Nodules in Micro-CT Imaging Using Deep Learning / M. D. Holbrook, D. P. Clark, R. Patel [et al.] // Tomography. – 2021. – Vol. 7, no. 1. – P. 32. – DOI: 10.3390/tomography7030032.

5. Zhang, H. Pulmonary nodules detection based on multi-scale attention networks / H. Zhang, Y. Peng, Y. Guo // Scientific Reports. – 2022. – Vol. 12. – Art. 1466. – DOI: 10.1038/s41598-022-05372-y.

6. Hierarchical approach for pulmonary-nodule identification from CT images using YOLO model and a 3D neural network classifier / Y. Ahmadyar, A. Kamali-Asl, H. Arabi [et al.] // Radiological Physics and Technology. – 2024. – Vol. 17, no. 1. – P. 124–134. – DOI: 10.1007/s12194-023-00756-9. PMID: 37980315.

7. Towards Reliable and Explainable AI Model for Solid Pulmonary Nodule Diagnosis / C. Wang, Y. Liu, F. Wang [et al.] // arXiv preprint. – 2022. – URL: https://arxiv.org/abs/2206.12345 (date of access: 05.02.2024).

8. Shape and margin-aware lung nodule classification in low-dose CT images via soft activation mapping / Y. Lei, Y. Tian, H. Shan [et al.] // Medical Image Analysis. – 2020. – Vol. 60. – Art. 101628. – DOI:10.1016/j.media.2019.101628.

9. Murphy, A. Windowing (CT) / A. Murphy, Y. Baba // Radiopaedia.org. – URL: https://radiopaedia.org/articles/windowing-ct?lang=us (date of access: 26.09.2024). – DOI: 10.53347/rID-52108.

10. Jocher, G. Ultralytics YOLOv8 / G. Jocher, A. Chaurasia, J. Qiu. – 2023. – URL: https://github.com/ultralytics/ultralytics (date of access: 26.09.2024).

УДК 004.932.2

Статистический анализ применения искусственного интеллекта для обнаружения таблиц в документах PDF

Р. О. Кохнович
ООО «Дуал Лаб Бел», Минск, Беларусь
E-mail: roma3122001@gmail.com

Введение

В эпоху цифровой трансформации огромные объемы информации до сих пор хранятся в статичных, хотя и повсеместно распространенных форматах, таких как PDF. Особую сложность для автоматической обработки представляют собой таблицы – ключевые носители структурированных данных, будь то финансовые отчеты, научные статьи или техническая документация. Таблицы содержат высокую плотность фактов, однако их автоматическое обнаружение в реальных документах осложняется разнообразием верстки, присутствием сканов, смешением декоративных и семантических элементов и неоднозначностью самого понятия «таблица» в контексте PDF. На помощь приходят технологии искусственного интеллекта, а именно компьютерное зрение и глубокое обучение, которые позволяют детектировать и распознавать таблицы, воспринимая документ как целостное изображение.

Обнаружение таблиц – критический компонент аудита доступности PDF в соответствии со стандартами вроде WCAG 2 и PDF/UA. Для пользователей экранных дикторов таблица доступна только в том случае, если она корректно размечена в structure tree, содержит семантические элементы Table, TR, TH/TD, имеет связность заголовков с ячейками и не нарушает логический порядок чтения [1, 2].

Цель исследования – провести масштабную и объективную оценку качества работы современной модели Docling [3] в задаче детекции таблиц, чтобы получить статистически значимые результаты и выявить не только метрики, но и глубинные причины возможных ошибок.

Алгоритм работы

Работа была произведена с большим неоднородным датасетом реальных PDF. В качестве эталона была использована структура документа PDF, так называемое structure tree, где находится информация об элементах,

которые находятся в документе [4]. Важное допущение: structure tree отражает намерение создателя документа и не всегда совпадает с тем, что человек воспринимает как таблицу на странице. Единицей оценки был объект Table. Предсказания Docling сопоставлялись с эталонными элементами по метрике IoU, после чего подсчитывались true positives, false positives и false negatives. Масштаб датасета делает полную ручную проверку невыполнимой, а распараллеливание обработки страниц и документов в AWS обеспечивает приемлемое время отклика и стоимость. Обработка выполнялась пакетно в AWS: документы и промежуточные результаты хранились в S3, вычисления производились на EC2, а для контроля качества сохранялись изображения страниц с наложенными рамками эталона и предсказаний [5].

В результате проверки около 440 тыс. документов PDF были получены следующие результаты:

true positives – 835 586;

false positives – 632 074;

false negatives – 397 907.

Рассчитанные на их основе метрики: precision 56,9 %, recall 67,7 %, f1-мера 61,8 %.

На практике это означает, что модель находит примерно две из трех таблиц, присутствующих в структуре документа. Однако почти каждое второе срабатывание оказывается ошибочным или спорным относительно выбранного эталона. Это неудивительно: на реальных PDF граница между таблицей, списком, оглавлением и оформленным текстовым блоком часто размыта, а дизайнерские решения вроде рамок и заливок визуально имитируют таблицу без наличия ячеистой структуры.

Анализ ошибок

Для отделения собственно ошибки модели от неоднозначностей разметки была проведена ручная проверка 500 случайных таблиц. Результаты показали, что в около 77 % случаев Docling был прав в обнаружении таблиц, в 13 % – не прав, еще 10 % были отнесены к неопределенным случаям. Если исключить неопределенные примеры из подсчета, доля корректных среди определимых случаев достигает примерно 85,5 %. Этот эксперимент показывает, что заметная часть расхождений с эталоном объясняется именно разночтениями: то, что модель помечает как таблицу, человек нередко тоже склонен считать таблицей, но соответствующий элемент может отсутствовать в structure tree и наоборот. Далее представлены типичные примеры случаев true positives (рис. 1), false positives (рис. 2) и false negatives (рис. 3). Красным цветом отмечены таблицы, которые нашел Docling, синим – которые были в structure tree.

	長野リーグ（女子）	長野リーグ（男子）	1部リーグ（男子）	2部リーグ（男子）	3部リーグ（男子）	3部予選（男子）
構成	3～チーム	8チーム	1組8/2組8	1組6/2組6	1組6/2組6	未定
日程	3から4月	11から4月	11から4月	春3月秋11月	春3月秋11月	春4月秋11月
試合形式（大会規則）	リーグ戦	リーグ戦	リーグ戦 順位決定戦	リーグ戦	リーグ戦	リーグor ﾄｰﾅﾒﾝﾄ
昇格	1位北信越	1位北信越	1・2・3位 入れ替え戦	春：各組2位 秋:なし	春：各組2位 秋：各組2位	春：上位4ﾁｰﾑ 秋：上位4ﾁｰﾑ
降格		6・7・8位 入れ替え戦	各組7・8位	各組5・6位	各組5・6位	
試合方法（大会規則）	S3・D2 1セット	S2・D1 S３セット D３セット(10P)	S3・D2 1セット	S2・D1 1セット	S2・D1 1セット	S2・D1 1セット
オーダー制限	制限なし	暫定制限なし	制限なし	制限なし	制限なし	制限なし
参加費	¥18500 (3ﾁｰﾑの場合¥17000)	¥61500	¥61000	春¥21000 秋¥21000	春¥21000 秋¥21000	春¥16000 秋¥16000
試合前チーム練習（大会規則）	5分 １２分	5分 １２分	5分 １２分	サービス４本 １２分	サービス４本 １２分	サービス４本 １２分

Рис. 1. Пример true positive

En combinant les approches biologiques et physiques, reliant observations naturalistes, analyse des comportements de vol et modélisation aérodynamique, cette étude met en évidence le rôle de la morphologie dans l'adaptation à un nouveau milieu – la canopée. Elle représente une étape majeure dans la compréhension de l'évolution des niches écologiques de la forêt Amazonienne et du rôle de la coévolution entre forme des ailes et comportement de vol dans l'adaptation des papillons à des environnements contrastés.

Références :

Le Roy C., D. Amadori, S. Charberet, J. Windt, F. T. Muijres, V. Llaurens* and V. Debat*. Adaptive evolution of flight in Morpho butterflies. ***Science***, Novembre 2021. * equal contribution. DOI: 10.1126/science.abh2620

CONTACTS PRESSE
Muséum national d'Histoire naturelle
PRESSE@MNHN.FR
SAMYA RAMDANE : 01 40 79 54 40
BLANDINE PRIOUR : 01 40 79 53 87
SOPHIE MINODIER : 01 40 79 38 00
MNHN.FR

Рис. 2. Пример false positive

Pearson has *a net profit margin* of 9.8%, which is behind only Stride with 10%. The rest of the comparable companies have a mean of only -17.7, proving that Pearson is one of the best in the industry at streamlining its revenue. One reason for this could be that Pearson is one of only 4 exam boards in the whole of England, meaning there is less competition. This allows the firms to raise their prices and increase their profit margins. The firm also has an *asset turnover ratio* of 1.2x, which is higher than the mean of the comparable companies, suggesting that Pearsons's assets are productive. This is likely due to the fact that Pearson sells a lot of textbooks and exam papers, meaning they have high sales alongside a smaller asset base. Finally, *the financial leverage* of the company is measured by a debt/capital ratio of 30.3%. This is the 2nd lowest out of its peers, which is attractive because it shows Pearson has the cash flow to be able to fund new investments If opportunities arise, and also lowers the irks of the investment.

- From this, we can see that Pearsons strengths lie in its revenue generation, suggesting that the firm has good financial management and efficient production processes. However, the lower amount of debt means that Pearson may not be utilising its strong cash flow to the best of its abilities.

Company Name	TEV/ LTM Total Revenues	TEV/Gross Profit	TEV/ LTM EBITDA	TEV/ LTM EBIT	TEV/CF from Ops.	Price/BV 17/10/2024	TEV/ NTM Total Revenue	TEV/ NTM EBITDA	PEG Ratio NTM	Forward P/E NTM
High	3.5x	15.2x	26.7x	318.6x	29.9x	5.95x	3.3x	20.7x	0.2x	41.3x
Low	0.6x	0.8x	4.9x	10.1x	2.3x	0.43x	0.7x	2.8x	0.1x	1.9x
Mean	1.4x	4.8x	12.5x	95.0x	10.5x	2.38x	1.4x	10.3x	0.1x	19.8x
Median	1.0x	3.3x	9.2x	25.6x	6.0x	2.08x	1.0x	6.1x	0.1x	12.8x
Pearson plc	2.3x	4.6x	13.4x	15.7x	14.1x	1.9x	2.3x	10.1x	2.5x	16.0x
AcadeMedia AB (publ)	1.0x	3.4x	9.8x	12.0x	5.2x	1.1x	1.0x	4.5x	NA	7.7x
Chegg, Inc.	0.6x	0.8x	4.9x	318.6x	2.3x	0.4x	0.7x	2.8x	0.1x	1.9x
Coursera, Inc.	0.7x	1.4x	NM	NM	6.0x	2.0x	0.7x	17.5x	0.2x	41.3x
Bright Horizons Family Solutions Inc.	3.5x	15.2x	26.7x	39.1x	29.9x	6.0x	3.3x	20.7x	NA	35.2x
Stride, Inc.	1.2x	3.3x	8.6x	10.1x	9.1x	2.4x	1.2x	6.1x	NA	12.8x

Ratio analysis

There are some standout figures in the table above. The most relevant is Pearsons' forward P/E ratio of 16x, which suggests the company is relatively undervalued. A lower forward P/E indicates that analysts expect future earnings to increase, making Pearson an attractive investment.

Рис. 3. Пример false negative

Анализ ошибок позволил выявить несколько устойчивых сценариев, где ответы модели не совпадают с эталоном.

Случаи false negatives:

– реальные таблицы, которые отмечены в structure tree как изображения;

– многоколонная композиция и регулярные отступы обманывают модель, например списки;

– элементы дизайна с рамками и заливками.

Случаи false positives:

– таблицы без линий, построенные исключительно на выравнивании текста;

– сложные структуры, например объединенные ячейки и многоуровневые заголовки;

– модель может воспринять продолжение таблицы на следующей странице как отдельный объект.

Результаты после фильтрации

Была также произведена фильтрация используемого датасета, в ходе которой были убраны файлы с неправильными таблицами, таблицы с неверным синтаксисом structure tree и неправильным размещением ячеек. После чего на полученном датасете из 310 тыс. документов снова был запущен Docling. Результаты были следующие:

true positives – 693 149;

false positives – 272 216;

false negatives – 287 445.

Рассчитанные на их основе метрики: precision 71,8 %, recall 70,6 %, f1-мера 71,2 %.

При этом результаты ручной проверки еще 500 случайных таблиц показали, что модель все также верна примерно в 80 % случаев. Другими словами, значительная часть расхождений с эталоном объясняется не ошибками модели, а фундаментальным разночтением между structure tree и визуальным восприятием человека.

Заключение

Данное исследование показывает, что даже в крупных реальных корпусах существенная доля «табличных» визуальных объектов либо отсутствует в structure tree, либо размечена неоднозначно. Это означает, что опирающиеся на структуру автоматические проверки доступности не будут находить таблицы там, где они видны на странице, но не описаны в структуре, и могут находить там, где декоративные блоки ошибочно приняты за таблицы.

Несмотря на скромные метрики против structure tree анализ показывает, что искусственный интеллект, в частности модель Docling, является практичным инструментом для обнаружения таблиц в промышленных сценариях. Главное ограничение исследования состоит в том, что метрики привязаны к structure tree, которое лишь частично отражает визуальную семантику страниц. В результате одновременно измеряется качество модели и согласованность с авторской разметкой.

Перспективы улучшения лежат не в области фундаментальной переработки моделей, а в инженерной доработке контура их применения. Ключевыми направлениями могут быть улучшение обработки сканов и таблиц без линий, внедрение правил для фильтрации ложных срабатываний, а также активное обучение на спорных примерах, где мнение модели и эталона расходятся.

Список использованных источников

1. W3C. Web Content Accessibility Guidelines (WCAG) 2.2 / World Wide Web Consortium (W3C). – URL: https://www.w3.org/WAI/standards-guidelines/wcag (date of access: 24.09.2025).

2. ISO 14289-1:2014 Document management applications – Electronic document file format enhancement for accessibility – Part 1: Use of ISO 32000-1 (PDF/UA-1) / Международная организация по стандартизации (ISO). – 2014. – URL: https://www.iso.org/standard/64599.html (date of access: 24.09.2025).

3. Docling Documentation. – URL: https://github.com/docling-project/docling (date of access: 24.09.2025).

4. ISO 32000-2:2020. Document management – Portable document format – Part 2: PDF 2.0. – Международная организация по стандартизации, 2020.

5. Amazon Web Services (AWS). Official Documentation. – URL: https://docs.aws.amazon.com (date of access: 24.09.2025).

УДК 34.096

Принципы обработки персональных данных системами искусственного интеллекта: содержание и современное состояние правового регулирования

Я. А. Курто
Полоцкий государственный университет
имени Евфросинии Полоцкой, Новополоцк, Беларусь
E-mail: y.kurto@psu.by

Введение

В современную эпоху цифровой трансформации искусственный интеллект (ИИ) способствует изменению способов взаимодействия с информацией. Особую значимость приобрели вопросы правового регулирования в контексте защиты персональных данных при их обработке системами ИИ.

Актуальность исследования трансформации сферы ИИ подтверждается Концепцией национальной безопасности Республики Беларусь. Так, основные тенденции современного мира характеризуются «продвижением цифровых технологий, биотехнологий и искусственного интеллекта, которое формирует предпосылки для изменения структуры экономики, трансформации традиционных направлений внедрения производительных сил и финансовых потоков» [1]. В этом исследовании рассматриваются содержание принципов обработки персональных данных системами ИИ в контексте защиты персональных данных, а также современное состояние правового регулирования как в Беларуси, так и в зарубежных странах.

Основная часть

Понятие обработки персональных данных является основополагающим для рассмотрения вопросов охраны персональных данных при использовании технологии ИИ. Рассмотрим основные требования к данной процедуре, установленной законодателем, а также некоторые вопросы его практической реализации. Закон Республики Беларусь от 7 мая 2021 г. № 99-З «О защите персональных данных» (далее – Закон № 99-З) под обработкой персональных данных понимает любое действие или совокупность действий, совершаемые с персональными данными, включая сбор, систематизацию, хранение, изменение, использование, обезличивание, блокирование, распространение, предоставление и удаление [2].

В соответствии с Законом № 99-З оператором признается государственный орган, юридическое лицо Республики Беларусь, иная организация, физическое лицо, самостоятельно или совместно с иными указанными лицами организующие и (или) осуществляющие обработку персональных данных [2]. Возникает вопрос о том, кого следует признавать оператором персональных данных в случае использования технологии ИИ, потому как ИИ, хоть и есть технология, обучаемая и достигающая результатов на этой основе, однако таковым оператором не является. Данный вопрос поднимает иную тему обсуждения в правовом поле, а именно способность ИИ являться субъектом права. И. А. Филипова и В. Д. Коротеев отмечают «юридическую нецелесообразность признания за искусственным интеллектом статуса субъекта права в классическом понимании теории права» [3]. Таким образом, ИИ не может признаваться оператором персональных данных в связи с отсутствием правосубъектности и, следовательно, несоответствием признаку отношения к организации или физическому лицу согласно Закону №99-З. Предлагается признание разработчиков систем ИИ в качестве операторов персональных данных. Поэтому меры ответственности за нарушения требований к обработке персональных данных ИИ будут применяться в отношении лиц, уполномоченных на разработку систем ИИ, в рамках определенной организации.

Обработка персональных данных системами ИИ состоит из двух этапов: обучение – ИИ учится выявлять закономерности и выполнять требуемые задачи, использование – ИИ непосредственно применяет полученные данные для достижения поставленных целей. Входные данные, на основе которых ИИ проходит самообучение, могут состоять в том числе и из персональных данных.

Правовыми основаниями обработки являются согласие субъекта персональных данных, заключение или исполнение договора, обработка персональных данных, указанных в адресованном оператору документе, оформление и реализация трудовых (служебных) отношений, защита жизни, здоровья или иных жизненно важных интересов человека, обработка распространенных ранее персональных данных, выполнение обязанностей (полномочий), предусмотренных законодательными актами [2]. Отметим, что цели обработки персональных данных, в частности специальных, при которых согласие субъекта не требуется, установлены законодательством и расширительному толкованию не подлежат. В случае использования технологий ИИ наиболее вероятным основанием представляется обработка распространенных ранее персональных данных.

Независимо от правового основания обработки персональных данных операторы обязаны соблюдать общие требования к обработке персональных данных. Ведущим нормативным правовым актом, утверждающим

требования к обеспечению защиты персональных данных на территории Европейского союза, в том числе к обработке персональных данных, является «Общий регламент по защите персональных данных (GDPR)» (далее – Регламент), действующий на территории Европейского союза с 2018 г. [4]. Принципы обработки персональных данных установлены 5-й статьей Регламента, а именно принципы законности, справедливости и прозрачности, ограничения целью, минимизации данных, точности, ограничения хранения, целостности и конфиденциальности [4].

В Регламенте установлены нормы, которые нашли отражение в Законе № 99-З [2, 4]. В названном Законе определены как общие критерии по отношению к защите персональных данных, так и требования к разработчикам ИИ при обработке системой персональных данных.

Закон № 99-З выделяет такие требования к обработке персональных данных, как законность, соразмерность, справедливость, ограничение цели их обработки, наличие правового основания обработки персональных данных, ограничение хранения персональных данных и прозрачный характер обработки персональных данных [2].

Законность предполагает выполнение оператором обязанностей, возложенных на него законом. По данному принципу напрямую реализуется и принцип верховенства права, согласно которому органы и должностные лица, организации и граждане действуют в пределах Конституции Республики Беларусь и принятых в соответствии с ней актов законодательства [2, 5]. Так, персональные данные человека могут быть размещены на общедоступных веб-сайтах. Генеративные ИИ собирают данные из веб-страниц для последующего обучения системы. Т. Куру отмечает, что «сканирование сайтов разработчиками генеративных ИИ будет признаваться соответствующим закону только после извлечения конфиденциальных данных из собираемой информации» [6].

Принципы соразмерности и справедливости обработки персональных данных нашли свое отражение в п. 2 ст. 4 Закона № 99-З, в соответствии с ним обработка должна являться соразмерной заявленным целям и обеспечивать на всех этапах обработки справедливое соотношение интересов всех заинтересованных лиц [2]. Следовательно, персональные данные должны быть обработаны лишь в том случае, если достижение цели и получение результатов деятельности ИИ без их обработки невозможны. Предполагается, что анонимизация данных позволит получить необходимый разработчикам ИИ результат, не подвергнет субъекты персональных данных угрозам и обеспечит защиту личной информации.

Требование ограничения цели обработки персональных данных устанавливается для недопущения превышения прав оператора и защиты прав и интересов субъектов персональных данных. Цель должна быть конкрет-

ной, законной, а также заявленной до обработки. Недопустима избыточность [2], персональные данные должны собираться для конкретных, отчетливых и законных целей и не обрабатываться в последующем. Принцип ограничения целью может быть нарушен в случае применения технологии ИИ, а нарушением будет считаться любое использование персональных данных ИИ в несоответствии заявленной цели. Например, при создании баз данных персональной информации, полученной ИИ, и (или) продаже данных третьим лицам.

Прозрачность заключается в возможности получения субъектом персональных данных информации, касающейся их обработки. Данный принцип представляется самым труднодостижимым, так как в большинстве случаев персональные данные аккумулируются в базы данных ИИ из общедоступных источников, и пользователи сети Интернет не догадываются о том, куда в итоге попадают их персональные данные. Более 92 % самых популярных веб-сайтов отслеживают пользователей без их уведомления, а в случае когда пользователи прямо отказывались от этого, то более 85 % опрошенных веб-сайтов все равно продолжали это делать [7]. Поэтому право субъекта персональных данных на получение информации, касающейся обработки персональных данных, является труднореализуемым.

И. И. Литвин считает, что «при обработке персональных данных ИИ сведения остаются между субъектом данных и автоматизированной системой, который действует в рамках своего алгоритма, и предлагает создать условия, при которых оператор данных не будет иметь прямого доступа к данным» [8, с. 115]. По мнению автора, данная концепция может повлечь за собой потенциальные правовые коллизии. Например, в случае если ИИ запомнит персональную информацию, они потенциально могут оказаться в общем доступе либо в случае кибератаки, либо при предоставлении такой информации самим ИИ вследствие ошибки. Необходимо наличие человеческого фактора, что позволит предотвратить потенциальные угрозы в отношении безопасности персональных данных. Вместе с тем отсутствие у оператора доступа к обрабатываемой информации усложнит реализацию принципа прозрачности обработки персональных данных.

Требование достоверности обрабатываемых данных направлено на недопущение обработки устаревшей или недействительной информации. Оператор обязан принимать меры по обеспечению достоверности обрабатываемых им персональных данных, при необходимости обновлять их [2]. Обработка действительной персональной информации является выгодной для разработчиков ИИ, так как от этого зависит эффективность системы ИИ. Однако обработка актуальных персональных данных должна также являться целью, к которой разработчики ИИ будут стремиться в исполнение законодательства, а не из личных целей.

Обязательство к ограничению хранения персональных данных установлено с целью снижения рисков по отношению к персональным данным и защиты прав и законных интересов субъектов [2]. В случае если такая информация будет храниться разработчиками ИИ свыше установленного законодательством срока и не соответствовать первоначальной цели обработки персональных данных, речь будет идти о наличии угрозы незаконного использования информации. В соответствии с этим принципом персональные данные должны храниться в форме, позволяющей идентифицировать субъектов данных не дольше, чем это необходимо для целей, для которых эти данные обрабатываются.

Принцип ограничения хранения персональных данных напрямую связан с принципом минимизации данных. Эффективность ИИ зависит от количества информации, получаемой на этапе аннотирования данных, следовательно, чем больше информации получит ИИ, тем результативнее будет работать система.

Принятие изменений в Закон № 99-З, а именно введение принципа минимизации данных, установленного Регламентом, позволит сдерживать стремление к получению избыточного количества персональных данных, ограничивать обработку данных задачами системы ИИ и необходимым объемом информации для их выполнения. В частности, для достижения законной цели обработки информации зачастую достаточно лишь полностью анонимных данных.

Однако не во всех случаях разработчики ИИ безоговорочно соблюдают установленные требования. Типичными нарушениями являются несоблюдение общих требований к обработке персональных данных, обработка персональных данных без наличия правовых оснований либо ненадлежащее правовое основание для этого и неисполнение данных условий, предъявляемых к получению согласия.

В 2025 г. компания Luka Inc., создатель чат-бота Replika, была оштрафована Итальянским органом по защите данных на 5 млн евро за нарушения регламента на основании обработки данных без правовых оснований, а также нарушения принципов прозрачности и целевого ограничения обработки персональных данных [9].

Обратим внимание на правовое регулирование вопроса в зарубежных странах. Среди иных зарубежных правовых актов, регулирующих требования к операторам при обработке персональных данных, стоит выделить «AI Act» (Европейский союз), Закон о кибербезопасности (the Cybersecurity Law of the People's Republic of China, CSL) и Закон о безопасности данных (the Data Security Law, DSL) (Китайская Народная Республика), Указ Президента Российской Федерации от 10 октября 2019 г. № 490 «О развитии

искусственного интеллекта в Российской Федерации» и Федеральный закон от 27.07.2006 N 152-ФЗ «О персональных данных».

Национальное правовое регулирование ИИ в сфере защиты персональных данных характеризуется менее развитой нормативно-правовой базой по сравнению с Китаем и Россией. Данная ситуация обусловлена стремительным прогрессом ИИ-технологий, который опережает возможности законодательной системы по адекватному реагированию на возникающие правовые вызовы.

В Беларуси регулирование сферы обработки персональных данных ограничивается общими положениями Закона № 99-З, который не содержит специализированных требований для обработки персональных данных ИИ-системами. Правовой статус ИИ-технологий остается неопределенным, отсутствуют специальные процедуры получения согласия субъектов данных для обработки их информации ИИ-системами. Предлагается установление требования к обработке персональных данных с использованием ИИ только при условии соблюдения общих требований Закона № 99-З, в том числе получения письменного согласия субъекта персональных данных, либо обязательного обезличивания персональных данных.

В 2025 г. был принят модельный закон «О технологиях искусственного интеллекта» [10], который направлен на принятие единой терминологии и принципов использования ИИ в странах СНГ. Названным законом регулируются «общественные отношения, возникающие на всех стадиях жизненного цикла технологий искусственного интеллекта: исследования, проектирования, разработки, оценки и проверки, эксплуатации и обслуживания, мониторинга и контроля, утилизации и иных, предусмотренных национальным законодательством». Более того, модельный закон устанавливает принцип защиты персональных данных, согласно которому «при использовании технологий искусственного интеллекта персональные данные подлежат защите в соответствии с национальным законодательством о защите персональных данных» [10].

Заключение

На сегодняшний день в Беларуси существует системный разрыв между стремительным развитием ИИ-технологий и существующим правовым регулированием в сфере защиты персональных данных. Принятие в 2025 г. модельного закона создает методологическую основу для унификации подходов к регулированию ИИ и потенциально может способствовать более полному регулированию вопроса обработки персональных данных системами ИИ. Предлагается использовать международный опыт правового регулирования исследуемой сферы для создания современной, эффектив-

ной системы правового регулирования обработки персональных данных системами ИИ, а также общие требования, установленные Законом № 99-З.

Список использованных источников

1. Концепция национальной безопасности Республики Беларусь : решение Всебелорус. нар. собр., 25 апр. 2024 г., № 5. – Минск : Нац. центр правовой информ. Респ. Беларусь, 2024. – 62 с.

2. О защите персональных данных : Закон Респ. Беларусь от 7 мая 2021 г. № 99-З // ЭТАЛОН : информ.-поисковая система (дата обращения: 02.10.2025).

3. Филипова, И. А. Будущее искусственного интеллекта: объект или субъект права? / И. А. Филипова, В. Д. Коротеев // Journal of Digital Technologies and Law. – 2023. – № 1 (2). – С. 359–385.

4. General Data Protection Regulation // EUR-Lex. – URL: https://eur-lex.europa.eu/eli/reg/2016/679/oj (date of access: 05.10.2024).

5. Конституция (Основной Закон) Белорусской Советской Социалистической Республики : принята на внеочередной девятой сессии Верховного Совета БССР девятого созыва 14 апреля 1978 года // Национальный правовой Интернет-портал Республики Беларусь. – URL: https://pravo.by/pravovaya-informatsiya/pomniki-gistoryi-prava-belarusi/kanstytutsyynae-prava-belarusi/kanstytutsyi-belarusi/konstitutsiya-1978-goda-s-izmeneniyami-i-dopolneniyami/ (дата обращения: 23.04.2025).

6. Kuru, T. Lawfulness of the mass processing of publicly accessible online data to train large language models / T. Kuru // International Data Privacy Law. – 2024. – Vol. 14, no. 4. – P. 326–345.

7. Lancieri, F. Narrowing Data Protection's Enforcement Gap / F. Lancieri // Maine Law Review. – Vol. 74, no. 1. – P. 17–18.

8. Литвин, И. И. Особенности сбора, обработки и защиты персональных данных искусственным интеллектом / И. И. Литвин // Вестник Уральского юридического института МВД России. – 2021. – № 4. – С. 112–118.

9. AI: the Italian Supervisory Authority fines company behind chatbot Replika // European Data Protection Board. – URL: https://www.edpb.europa.eu/news/national-news/2025/ai-italian-supervisory-authority-fines-company-behind-chatbot-replika_en (date of access: 04.10.2025).

10. Модельный закон «О технологиях искусственного интеллекта» / ОИПИ НАН Беларуси. – URL: https://uiip.by/doc/pdf/model_law.pdf (дата обращения: 04.10.2025).

УДК 004.8

Роль искусственного интеллекта в юриспруденции

А. А. Лабоцкая
Центр системного анализа и стратегических исследований НАН Беларуси, Минск
E-mail: nastena_labockaya@mail.ru

Введение

Сегодня искусственный интеллект (ИИ) стремительно проникает во все сферы деятельности, в том числе и юриспруденцию. Благодаря машинному обучению, обработке естественного языка и анализу больших данных ИИ открывает новые возможности для применения в правовой практике.

В юриспруденции ИИ используется в разных масштабах и направлениях в зависимости от конкретной страны. Например, в США юридическая фирма Baker & Hostetler подключила к бизнес-процессам робота-юриста, который работает на основе ИИ. Росс, разработанный IBM на базе компьютера Watson, возглавляет отдел по делам банкротства. Другая команда американских программистов создала алгоритм, который может предсказывать верное решение Верховного суда США в 70 % случаев. Также в Великобритании 18-летний программист создал сайт donotpay.co.uk, который помогает автовладельцам автоматически составить и отправить апелляцию на полученную квитанцию о штрафе [1].

Понятие искусственного интеллекта и статистика его использования в юриспруденции

Искусственный интеллект (англ. artificial intelligence, AI) – термин относится к области компьютерных технологий, занимающейся разработкой интеллектуальных машин, которые обладают схожим с человеческим мышлением, работают и реагируют как люди [2].

Вопросы, связанные с внедрением ИИ в юриспруденцию, на сегодняшний день вызывают повышенный интерес как у белорусских правоведов, судей и практикующих юристов, так и у зарубежных коллег.

В настоящее время юристы используют инструменты ИИ для просмотра документов (77 %), юридических исследований (74 %), обобщения документов (74 %), составления сводок или служебных записок (59 %) (рис. 1) [3].



Рис. 1. Использование ИИ юристами

Рассмотрим использование генеративного сервиса ChatGPT.

ChatGPT (англ. *generative pre-trained transformer* – генеративный предобученный трансформер) – чат-бот с генеративным ИИ, разработанный компанией OpenAI и способный работать в диалоговом режиме, поддерживающем запросы на естественных языках [4].

Образовательной платформой Moscow Digital School (входит в Ultimate Education) проведено исследование об использовании ChatGPT, которое показало, что 42,9 % опрошенных юристов уже его используют в работе, 25 % планируют начать, а остальные не собираются применять или не рассматривают эту возможность. Среди юристов, уже использующих ChatGPT (рис. 2), 43,8 % осуществляют поиск необходимой информации и ответов на вопросы, 19,2 % составляют типовые юридические документы, а 15,7 % выполняют подготовку научных диссертаций и монографий [5].



Рис. 2. Использование юристами ChatGPT

По мнению респондентов (рис. 3), наиболее полезное применение чат-ботов в автоматизации рутинных задач (73,6 %), подготовке типовых юридических документов (61,4 %), маркетинге и PR (52,6 %), научной деятельности (29,8 %) и подготовке стратегии судебной защиты (12,2 %). 8,7 % респондентов не видят пользы в использовании ИИ [5].

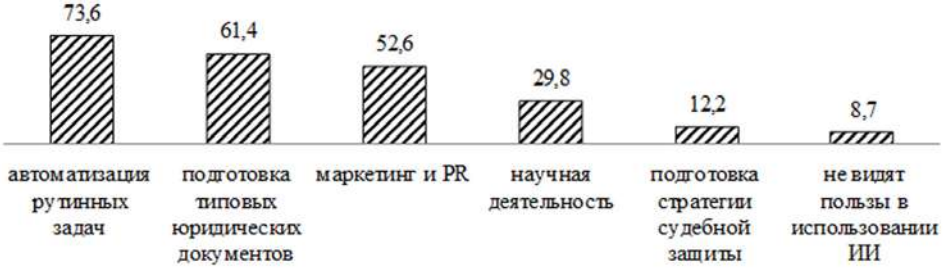


Рис. 3. Применение чат-ботов

Для многих практикующих адвокатов и юристов ИИ стал технологическим прорывом (табл. 1, 2).

Таблица 1

Использование ИИ в юриспруденции [6]

Наименование	Содержание
Анализ документов	ИИ способен за считанные минуты проверять договоры и другие тексты на соответствие законодательству, находить риски и противоречия. Например, нейросеть может проанализировать сложный контракт объемом в 50 страниц и выделить проблемные пункты гораздо быстрее человека. Также существуют решения для распознавания скан-копий документов и сравнения разных версий договоров, что значительно ускоряет юридическую экспертизу
Поиск правовых норм и консультации	Чат-боты, обученные на базе законов и судебной практики, отвечают на вопросы в естественной форме. Например, на вопрос «какая ответственность грозит за нарушение договора поставки?» ИИ может дать развернутый ответ со ссылками на нормы гражданского кодекса. Такие системы экономят часы, которые раньше тратились на ручной поиск по нормативным правовым актам и базам данных
Генерация юридических документов	Продвинутые сервисы умеют автоматически составлять черновики исковых заявлений, жалоб и договоров. Пользователю достаточно ввести основные параметры (стороны, существенные условия), и генеративная модель на основе больших языковых сетей (GPT) сформирует шаблон документа. Адвокату, юристу останется только проверить и при необходимости скорректировать проект
Судебные аналитика и прогноз	ИИ может оценивать перспективы дела на основании статистики ранее вынесенных решений судов. Аналогично тому, как канадская платформа анализирует прецеденты, российские решения позволяют спрогнозировать вероятность удовлетворения иска, опираясь на судебную практику России. Такой функционал помогает адвокату оценить шансы доверителя и выбрать оптимальную стратегию

В Республике Беларусь примерами технологий ИИ в юридической сфере можно назвать следующие:

1) крупные компании и юристы создают программы, чат-ботов, где клиенты самостоятельно в режиме онлайн могут подготовить несложное исковое заявление, составить договор, зарегистрировать бизнес, задекларировать доходы или получить консультацию (онлайн-конструктор документов Rules.by);

2) юридические интернет-порталы, такие как Pravo.by и belnotary.by, объединяют широкую базу юристов, нотариусов, адвокатов, которые могут проконсультировать клиента по типовым вопросам в режиме онлайн. Главное преимущество – юридическую консультацию можно получить в сети Интернет, не теряя времени на посещение специалиста [8].

Отметим, что в рамках цифровизации государственного управления главой государства Республики Беларусь одобрен проект интеграции си-

стем ИИ в автоматизированную информационную систему «Нормотворчество», который оптимизирует процесс создания нормативно-правовых актов с помощью анализа существующей законодательной базы на предмет дублирования положений и оценки целесообразности принятия новых документов в предложенной редакции [9].

Таблица 2

Возможные уровни использования ИИ в судопроизводстве [7]

Уровень использования ИИ	Содержание
Техническая работа как помощника судьи	Судья лично исследует доказательства, имея под рукой результаты анализа, проведенного ИИ. При этом техническая работа ИИ не содержит прямых выводов по вопросам, входящим в предмет доказывания по конкретному делу. Например, ИИ может делать расшифровки (транскрипты) аудиозаписей, проверять наличие в материалах дела выписок из реестров в отношении юридических лиц, объектов права промышленной собственности или объектов недвижимости, в отношении которых заявлены иски. То есть ИИ анализирует доказательства и проверяет корректность представленных документов, а судья изучает имеющиеся доказательства, приведенные аналитикой ИИ, и самостоятельно делает выводы обо всех значимых для дела обстоятельствах
Оценка доказательств	Судья опирается на ту оценку доказательств, которую предоставил ИИ. В данном случае вывод относительно обстоятельств, имеющих значение для рассмотрения дела, делается ИИ, а не судьей. Необходимо обратить снимание на технологии использования ИИ в так называемом предиктивном (прогнозируемом) правосудии (justice predictive). По мнению доктора юридических наук, профессора, судьи Верховного Суда Российской Федерации В. В. Момотова технология прогнозируемого правосудия может рассматриваться только в качестве помощника, но не замены судьи. При использовании ИИ в предиктивном правосудии существует риск, что судья и участники процесса начнут зависеть от заранее сделанных прогнозов относительно итогов процесса. Судья в случае несогласия с предложенным решением ИИ будет вынужден не только обосновать, но и объяснить причину несогласия
Рассмотрение дела	ИИ самостоятельно принимает судебные акты. В таком случае, в отличие от ИИ, делающего выводы по существу спора, судья фактически не анализирует доказательства, имеющие значение для правильного разрешения дела

Заключение

Таким образом, ИИ оказывает значительное влияние на юридическую сферу, повышая эффективность и скорость обработки информации, сокращая расходы и помогая юристам принимать решения, а также улучшает доступ к правовой информации. ИИ упрощает работу юристов и специалистов в других областях, однако необходима тщательная проработка правовых аспектов, этических вопросов и механизмов контроля.

Список использованных источников

1. Инвестиции в судебные споры и боты-юристы: как развивается Legal tech в Беларуси // KV.by. High-Tech Club. – URL: https://www.kv.by/blog/users/iskinvest/1054059-investicii-v-sudebnye-spory-i-boty-yuristy-kak-razvivaetsya-legal-tech (дата обращения: 18.09.2025).

2. Искусственный интеллект // iot.ru. Новости Интернета вещей. – URL: https://iot.ru/wiki/iskusstvennyy-intellekt (дата обращения: 18.09.2025).

3. How AI is transforming the legal profession. – URL: https://legal.thomsonreuters.com/blog/how-ai-is-transforming-the-legal-profession/ (date of access: 18.09.2025).

4. ChatGPT // ВикипедиЯ. Свободная энциклопедия. – URL: https://ru.wikipedia.org/wiki/ChatGPT (дата обращения: 18.09.2025).

5. Более 40 процентов юристов уже используют искусственный интеллект в работе // RG.RU. Интернет-портал «Российской газеты». – URL: https://rg.ru/2023/07/15/bolee-40-procentov-iuristov-uzhe-ispolzuiut-iskusstvennyj-intellekt-v-rabote.html (дата обращения: 18.09.2025).

6. Искусственный интеллект в юридической сфере // advgazeta.ru. Адвокатская газета. Орган Федеральной палаты адвокатов РФ. – URL: https://www.advgazeta.ru/mneniya/iskusstvennyy-intellekt-v-yuridicheskoy-sfere/ (дата обращения: 22.09.2025).

7. Козорезова, Р. В. Использование искусственного интеллекта в судопроизводстве: реальность и перспективы / Р. В. Козорезова // Журнал «Судовы веснік». – 2023. – № 3. – URL: https://court.gov.by/ru/sup_court/int_prop/publ/b7ad4297f4a64efb.html (дата обращения: 22.09.2025).

8. Инвестиции в судебные споры и боты-юристы: как развивается Legal tech в Беларуси // KV.by. High-Tech Club. – URL: https://www.kv.by/blog/users/iskinvest/1054059-investicii-v-sudebnye-spory-i-boty-yuristy-kak-razvivaetsya-legal-tech (дата обращения: 29.09.2025).

9. Утвердили внедрение ИИ в разработку нормативно-правовых актов РБ // IXBT LIVE. – URL: https://www.ixbt.com/live/sw/lukashenko-utverdil-vnedrenie-iivrazrabotku-normativno-pravovyh-aktov-rb.html (дата обращения: 03.10.2025).

УДК 004.8

Цифровая трансформация государственных органов на базе искусственного интеллекта

А. И. Лапуцкий✉, **А. В. Сидорцов**
ООО «Программные решения плюс»,
Минск, Беларусь
E-mail: alaputski@sws.by, info@sws.by

Сложность и разобщенность ландшафта государственных информационных систем требуют подхода, объединяющего данные из различных систем с целью их анализа и реализации ранее недоступных управленческих функций. Предлагаемая платформа формирует единый контур приема, хранения, очистки, нормализации и анализа данных, а также исполнения и оптимизации процессов. В платформе реализован корпоративный искусственный интеллект, предоставляющий контекстно релевантные ответы и доступ к инструментам выполнения действий.

Список сокращений и специальных терминов:

ML (Machine Learning) – машинное обучение.

AI (Artificial Intelligence) – искусственный интеллект.

AI-агент – программное обеспечение, собирающее данные и автономно выполняющее бизнес-процессы на основе моделей искусственного интеллекта.

LLM (Large Language Model) – большая языковая модель, обученная на крупных корпусах текстов для генерации и понимания естественного языка.

Prompt (промпт) – задание в текстовой форме для LLM.

Embedding (эмбеддинг) – численное векторное представление текста (высказывания), учитывающее семантическую близость между высказываниями.

Vector Store (векторное хранилище) – база данных для хранения эмбеддингов, поддерживающая поиск по близости.

RAG (Retrieval-Augmented Generation) – архитектурный шаблон, при котором генерация ответа моделью обогащается релевантными документами, найденными в векторном хранилище.

MCP (Model Context Protocol) – интерфейс вызова инструментов и действий из AI-агента.

Оркестрация – процесс автоматизации управления, координации и организации работы компонентов программных систем.

BPMN (Business Process Model and Notation) – стандарт нотации для моделирования бизнес-процессов.

ETL/ELT (Extract Transform Load, Extract Load Transform) – процессы загрузки и преобразования данных, различающиеся порядком операций преобразования и загрузки данных.

Observability (наблюдаемость) – совокупность метрик, логов и трассировок для анализа состояния системы.

Low-code – подход к разработке, предполагающий визуальное моделирование и минимизацию процессов кодирования (программирования).

Данный список сформирован исходя из многолетней практики разработки программных систем для госорганов, куда вошли ключевые аспекты:

1. Необходимость интеграции и обработки данных разрозненных информационных систем для повышения эффективности управления и получения целостной аналитической картины, которая не может быть выведена из каждой системы в отдельности.

2. Необходимость оцифровки, анализа и классификации большого количества бумажных документов, в том числе архивных (находящихся на долгосрочном хранении). В данном аспекте требуется надежное распознавание, извлечение структурированных сущностей, тематическая классификация и дедупликация, сохранение юридической значимости документа.

3. Необходимость соединения рабочих процессов госоргана на базе одной платформы с целью повышения операционной эффективности и качества принятия управленческих решений.

4. Оптимизация процессов управления маршрутизацией документов при прохождении различных видов согласований, экспертизы, административных процедур посредством применения методов машинного обучения и формирования оптимальных процессов с помощью AI.

Предлагается комплексное решение, которое включает как программные модули, доказавшие свою эффективность в ходе промышленной эксплуатации, так и инновационные решения, связанные с корпоративным AI, в настоящее время проходящие тестирование реальными сценариями.

Инновационная архитектура, предлагаемая в качестве решения указанных проблем, включает в себя следующие компоненты (рис. 1):

– программное и инфраструктурное ядро распределенной отказоустойчивой системы;

– централизованное озеро данных, интегрирующее данные из различных систем госоргана, а также внешних источников информации, предоставляющих необходимые сведения для осуществления его деятельности;

– способы пополнения озера данными (подписка на API сторонних систем, прием данных потоковым или пакетным способом, добавление данных самими пользователями);

– способы очистки и трансформации данных, в том числе с использованием алгоритмов машинного обучения;

– агентный корпоративный AI, который работает с данными платформы и позволяет пользователям реализовывать собственные бизнес-сценарии.

– аналитическую платформу.

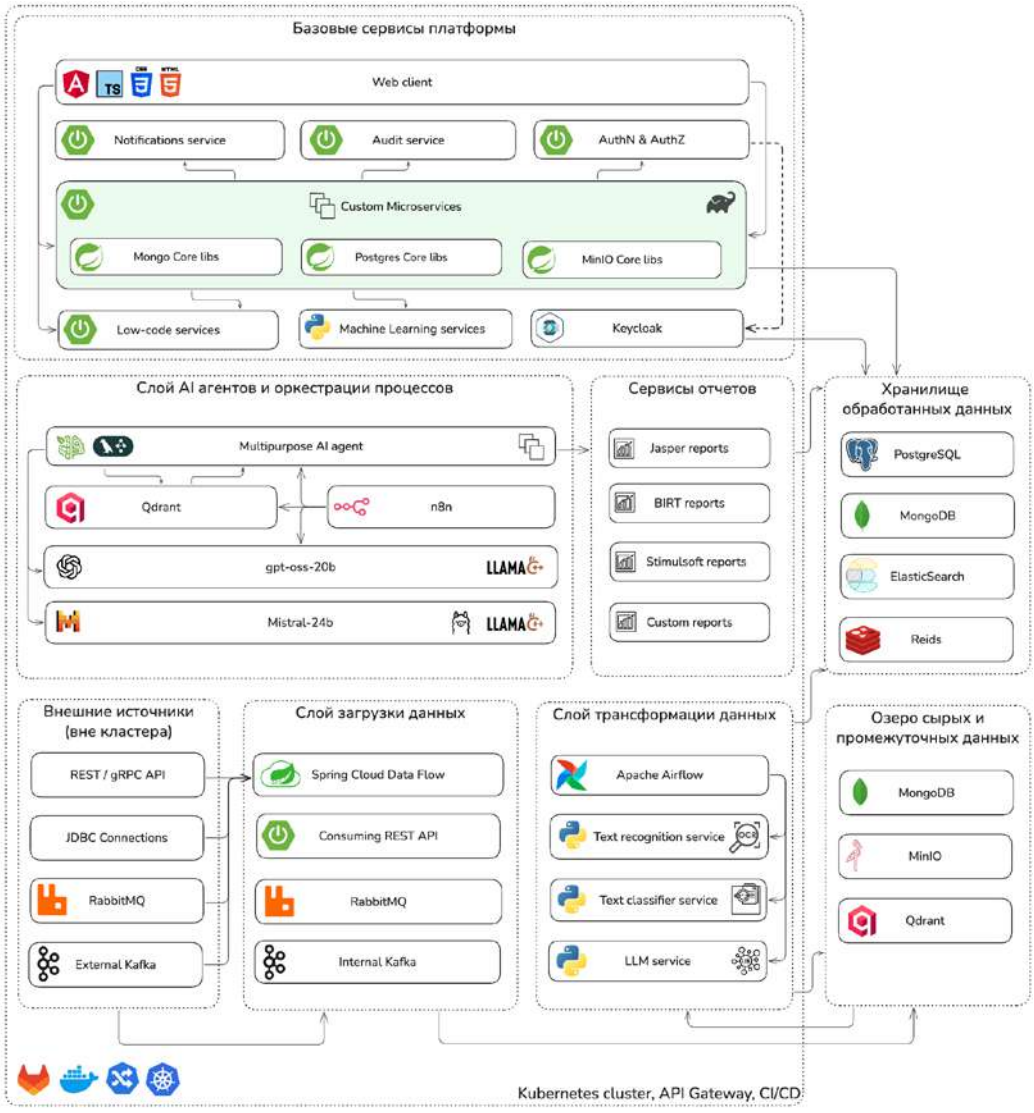


Рис. 1. Архитектура платформы с указанием используемых технологий

Инфраструктурное ядро реализовано как распределенная отказоустойчивая система, включающая в себя наборы взаимосвязанных инструментов, которые обеспечивают функционирование комплексов системного администрирования, защиты информации, аутентификации и авторизации, low-code-редактора по управлению структурами данных и многих других. Микросервисная архитектура обеспечивает горизонтальную масштабиру-

емость и изоляцию отказов [1–3]. Встроены Gitlab CI/CD, наблюдаемость (метрики, логи, трассировки).

Централизованное озеро данных объединяет внутренние и внешние источники, выделяя зоны сырых данных, очищенных представлений и потребительских витрин. Метаданные записываются в озеро данных параллельно с основным потоком данных и используются в low-code-редакторе и других функциональных частях системы.

Каналы ETL и ELT реализованы через подписку на API и шины событий, потоковую и пакетную загрузку, а также пользовательские загрузки с валидацией [4]. Очистка и трансформация охватывают нормализацию, обогащение, дедупликацию, извлечение признаков и ML-задачи классификации и регрессии. Особое внимание уделено проблеме регулярного нивелирования технического долга по данным, моделям и инфраструктуре [5]. В частности, загрузчики данных, программное ядро моделей машинного обучения и CI/CD используют Agile-подход к разработке в рамках жизненного цикла разработки конкретного проекта (SDLC).

Корпоративный AI в технической части опирается на LLM, векторное хранилище данных и оркестратор процессов n8n. Использование LLM позволяет значительно сократить время на обучение моделей для специфических бизнес-задач благодаря их способности к few-shot-обучению [6]. Few-shot-обучение – это подход, при котором модель может успешно выполнять задачи, имея лишь несколько примеров для обучения. Это позволяет моделям адаптироваться к новым задачам без необходимости обширного дополнительного обучения. В настоящий момент оркестратор процессов n8n имеет доступ к таким инструментам, как вызов API-микросервисов и отправка почтовых сообщений. Наличие цикла обратной связи позволяет администраторам системы с помощью AI-агентов вносить управляемые изменения в правила бизнес-процессов, обеспечивая эволюционное развитие.

Аналитическая подсистема предоставляет средства управления представлением данных в виде диаграмм, графов, средств формирования отчетов, дашбордов и OLAP-витрин данных с версионированием артефактов и соблюдением регламентов.

В целом архитектура соответствует 12-Factor-методологии (конфигурация через окружение, неизменяемые сборки, разделение build/release/run, логирование как поток событий и т. д.) [7], и The Reactive Manifesto (асинхронный обмен сообщениями, гибкость, устойчивость, отзывчивость) [8]. Для реализации приведенной архитектуры используются как фреймворки, являющиеся промышленными стандартами, так и инновационные решения.

Ядро платформы представляет собой микросервисную архитектуру на базе Kubernetes, Java, Spring Framework, TypeScript, Angular, Keycloak, Gradle and Docker. Также, используется фреймворк собственной разработ-

ки, который посредством специального конструктора (low-code-редактор структур) позволяет создавать и визуализировать структуры данных, интегрируя их в бизнес-процессы без дополнительного программирования.

Озеро данных основывается на использовании таких инструментов, как MinIO и MongoDB – для сырых, промежуточных данных и метаданных, PostgreSQL – для очищенных данных, Elasticsearch – для хранения логов и Redis – для кэширования. Пополнение озера данными осуществляется с помощью Kafka, RabbitMQ, Spring Cloud Data Flow, gRPC, REST API. Отдельным важнейшим каналом добавления данных является оцифровка документов, представленных в виде файлов различных форматов, в том числе графических (PDF, JPG, PNG).

Очистка и трансформация данных реализованы с использованием Python, PyTorch, scikit-learn, transformers [9], Bert, PyCaret, CatBoost, XGBoost, Optuna, а также Apache Airflow (для оркестрации процессов).

Корпоративный AI реализован на базе Spring AI, в качестве LLM используются Mistral-24b и gpt-oss-20b, llama.cpp – для запуска в облаке, Ollama – для локального запуска, Qdrant – для хранения эмбеддингов, Docker и Kubernetes. Также в текущее время реализуются рабочие прототипы на базе технологий LangGraph, n8n и pgvector для сценариев с гибридным поиском.

Аналитическая платформа использует Stimulsoft Reports, Stimulsoft Dashboards, JasperRepotrs, BIRT и собственные разработки.

Рассмотрим сценарий использования AI в платформе (рис. 2). Процесс начинается с централизованного поступления данных в озеро с параллельной индексацией содержимого в векторное хранилище. Индексация учитывает такие типы источников, как тексты, таблицы и смешанные документы.

Пользователь формулирует промпты (запросы) к LLM с указанием бизнес-сценариев, указывая область рассматриваемых документов прямо в интерфейсе пользователя на клиенте. Агент выполняет поиск в векторном хранилище, агрегирует найденные фрагменты и генерирует результат в заданном формате. Далее агент через MCP-интерфейсы способен использовать инструменты системы, а именно вызывать REST API микросервисов, сохранять результат в базу данных, отправлять отчет по электронной почте, инициировать следующую стадию процесса либо передавать управление агенту с иными инструкциями в качестве мастер-промпта. Данный подход открывает путь к автоматической модификации системы на основе комплекса знаний о ее состоянии.

На текущий момент в качестве экспериментального инструмента автоматизации рабочих процессов авторами используется n8n. Пользователем заводится схема n8n, которая реализует некий бизнес-сценарий. Специальный агент по работе с оркестрацией процессов предлагает корректировки схемы n8n (изменение порядка шагов, параметров и маршрутиза-

ции), которые затем валидируются администратором системы. Так формируется управляемый цикл обратной связи, при котором система эволюционирует на основе анализа данных и рекомендаций AI.

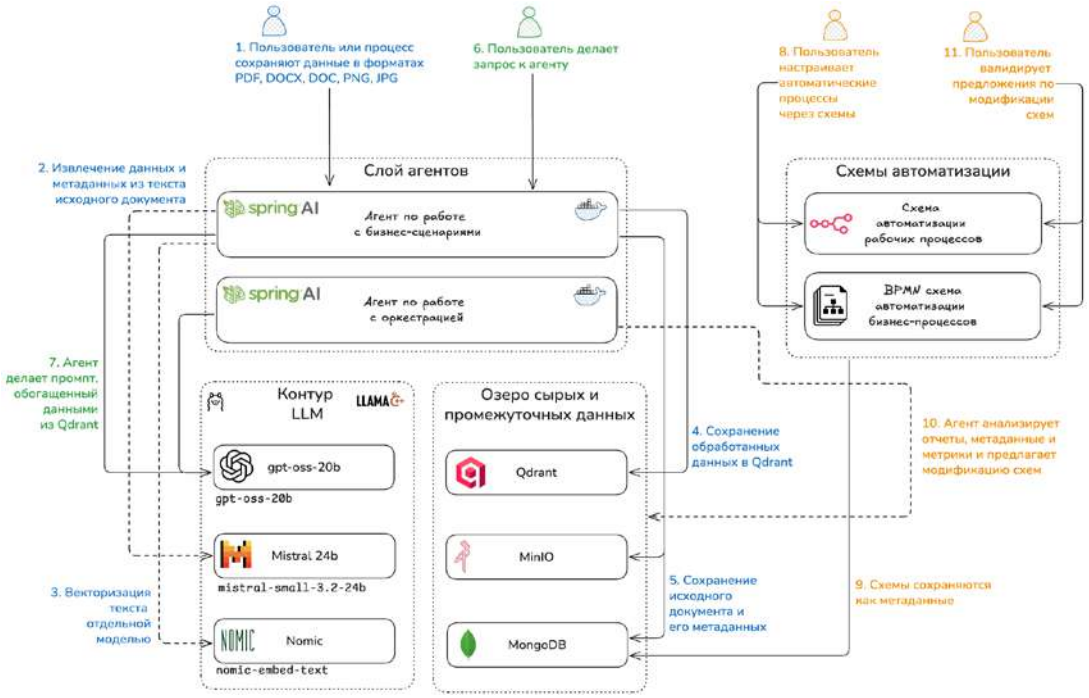


Рис. 2. Архитектура AI-агентов и цикл обратной связи

Рассмотрим контрольный пример, который реализует сценарий работы с данными о перевозке грузов. Корпус тестовых данных включает в себя по одной тысяче документов трех типов: заявки на перевозку грузов, акты выполненных работ и счета-фактуры.

Бизнес-процесс реализован в виде BPMN-схемы (заявка, подбор перевозчика, перевозка, акт выполненных работ, счет-фактура) и может быть интегрирован в платформу, например таким фреймворком, как Camunda. Также реализован ряд схем n8n с участием AI, которые анализируют заявки на перевозку и инициируют сотрудничество, рассчитывают котировки по данным о грузе и актуальным тарифам, рассылают в социальные сети сообщения со статусом перевозки, генерируют документы о перевозке и рассылают контрагентам. Здесь BPMN выступает как оркестратор всего процесса, а n8n – как исполнитель конкретных технических действий, которые могут инициироваться как BPMN-обработчиками, так и внешними акторами (заказчиком, перевозчиком, просто по расписанию) (рис. 3). Агент по работе с бизнес-сценариями получает от пользователя промпт на анализ трех типов документов и формирование стратегии оптимизации издержек. Результат должен быть сохранен в виде отчета. Пользователь работает с данным агентом пока, не получит приемлемую для себя стратегию.

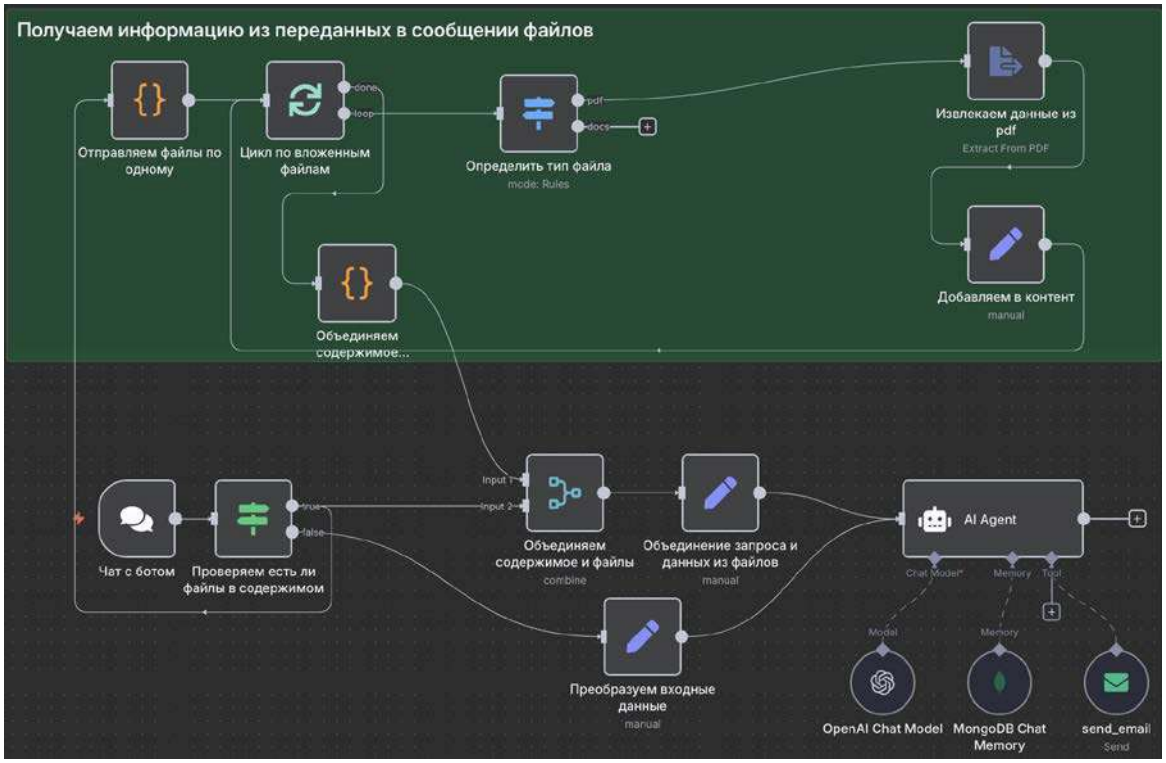


Рис. 3. Пример схемы n8n с анализом PDF, созданием отчета через AI-агента и отсылкой на e-mail пользователя

Агент по работе с оркестрацией анализирует поведение заказчиков (переписки, отзывы о работе), инциденты при перевозках, эффективность каждого этапа перевозки по конкретному маршруту, отказоустойчивость интеграций n8n и формирует предложения по обновлению BPMN- и n8n-схем. Результат включает отчет в виде текста, новые версии .bpmn и .json-файлов.

Рассмотрены актуальные проблемы госорганов и предложен способ их комплексного разрешения посредством создания программной платформы с использованием искусственного интеллекта. Данная платформа обеспечивает целостную интеграцию источников данных, оцифровку и интеллектуальную обработку документов, управление бизнес-процессами на основе AI. Архитектура опирается на проверенные практики проектирования распределенных систем и поддерживает управляемый цикл обратной связи, позволяющий планомерно улучшать данные и оркестрацию процессов. Данная разработка создает технологическую основу для регулируемой цифровой трансформации государственных органов в Республике Беларусь.

Список использованных источников

1. Bass, L. Software Architecture in Practice / L. Bass, P. Clements, R. Kazman. – 4th ed. – Boston : Addison-Wesley, 2021. – 624 p.

2. Tanenbaum, A. S. Distributed Systems: Principles and Paradigms / A. S. Tanenbaum, M. Steen. – 2nd ed. – Upper Saddle River, NJ : Pearson; Prentice Hall, 2007. – 765 p.

3. Newman, S. Monolith to Microservices: Evolutionary Patterns to Transform Your Monolith / S. Newman. – Sebastopol, CA : O'Reilly Media, 2019. – 272 p.

4. Kleppmann, M. Designing Data-Intensive Applications: The Big Ideas Behind Reliable, Scalable, and Maintainable Systems / M. Kleppmann. – Beijing ; Boston ; Farnham; Sebastopol ; Tokyo : O'Reilly Media, 2017. – 616 p.

5. Hidden Technical Debt in Machine Learning Systems / D. Sculley, G. Holt, D. Golovin [et al.] // Advances in Neural Information Processing Systems. – 2015. – Vol. 28. – P. 2503–2511.

6. Language Models are Few-Shot Learners / T. B. Brown, B. Mann, N. Ryder [et al.] // Advances in Neural Information Processing Systems. – 2020. – Vol. 33. – P. 1877–1901.

7. Wiggins, A. The Twelve-Factor App / A. Wiggins. – URL: https://12factor.net (date of access: 01.10.2025).

8. The Reactive Manifesto / J. Bonér, D. Farley, R. Kuhn, M. Thompson. – URL: https://www.reactivemanifesto.org (date of access: 01.10.2025).

9. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar [et al.] // Advances in Neural Information Processing Systems. – 2017. – Vol. 30. – P. 5998–6008.

УДК 004.8

Этические аспекты использования агентов искусственного интеллекта

А. И. Лойко
Белорусский национальный
технический университет, Минск

Введение

Технологии искусственного интеллекта (ИИ) благодаря разработчикам в настоящий момент представлены агентами ИИ. Это автономные программные системы ИИ, наделенные функциями имитации автономного мышления, итеративного планирования и автономного действия. Агенты ИИ обладают высокой степенью автономности, что проявляется в функции самостоятельно выполнять задачи для достижения цели, а также в функции к обучению и адаптации с течением времени в различных контекстах. Агенты ИИ смогут автоматизировать сложные задачи и совершенствовать принятие обоснованных решений. Они трансформируют здравоохранение, финансы и государственное управление. За этой трансформацией последуют социальные и правовые последствия, поскольку существующие структуры управления не учитывают особенности автономных агентов ИИ.

Возникает вопрос об этических аспектах использования агентов ИИ. Он может формулироваться в контексте трансформации существующих систем управления, но он также формулируется и в контексте применения концепта этоса к агентам ИИ. Когда речь идет о трансформации существующих структур управления, то фактически поднимается тема этики деловых и профессиональных отношений, в частности корпоративной этики. Она стала включать требования безопасности, конфиденциальности и коммерческой тайны. Ее категориальные структуры формируются концепциями прагматизма, утилитаризма, деонтологии и аксиологии. Этот концептуальный каркас на прикладном уровне представлен корпоративным этосом и миссией компании или банковской структуры.

Когда программисты стали важной частью деловой сферы, то были разработаны принципы этики программной инженерии. Необходимость их разработки была обусловлена феноменом социальной инженерии, которая создала систему практик и технологий манипулирования и хакинга корпоративных структур и пользователей. Вследствие этого компании стали пользоваться услугами этического хакинга. Этот хакинг осуществляют про-

фессиональные программисты, которые в рамках заказа тестируют корпоративную сеть на наличие у нее уязвимых мест и делают подробный отчет об обнаруженных уязвимостях с рекомендацией, как усилить структуры информационной безопасности.

Этические нормы и агенты ИИ

Более интересным представляется вопрос о возможности применения этических норм к агентам ИИ. В данном контексте предлагается разработать этос агента ИИ, сформулировать в его содержании руководящие принципы и исследовать моральную природу агентов ИИ [1]. Руководящие принципы предлагается сформулировать в виде модели этической технологии системы надзора. Для этого предполагается разработать онтологию новой предметной области на основе структур права и этической ответственности. Исходя из этого, агенты ИИ классифицируются по степени рисков для общества и корпоративной деятельности.

Методологическая нормативная модель закладывает основу для согласования технологий ИИ с общественными ценностями и обеспечения их ответственной разработки и внедрения. Она предлагает сбалансированную стратегию регулирования агентов ИИ. В основном эта методологическая модель рассчитана на разработчиков агентов ИИ. Проблема разработчиков заключается в том, что создаваемые ими когнитивные интеллектуальные структуры ИИ демонстрируют в процессе их практического использования феномены когнитивных искажений, галлюцинаций и предвзятости.

Получается, что разработчики неявно закладывают в семантику генеративного ИИ собственные когнитивные искажения и предвзятость. Компании уже столкнулись с судебными издержками по причине неадекватного поведения агентов ИИ при отборе персонала на свободные вакансии. В данном контексте возникает вопрос о качестве программной части технологий ИИ. От решения этой проблемы зависит перспектива роста инвестиций в технологии ИИ.

Вследствие обнаруженной практической проблемы возникает вопрос о том, что представляют собой агенты ИИ и какова их сущность. Отвечая на этот вопрос, обычно пишут, что агенты ИИ являются автономными сущностями, способными действовать самостоятельно, используя инструменты для достижения целей. Автономность обычно связывается с рациональностью, совместимой с этическими принципами. Без такого основания автономные системы ИИ могут генерировать цели способами, которые приводят к непредвиденным и потенциально опасным последствиям.

Рациональность агентов ИИ закладывается аппаратом четких и нечетких математических логик. Этот аппарат позволяет агентам ИИ принимать высокоуровневые решения и предпринимать действия, максимизирующие

производительность посредством данных и анализа данных. В философии агенты ИИ называются агентами симуляции убеждения, желания и намерения [2]. Эта рациональность формируется контекстуальной структурой и набором данных, в рамках которых агент ИИ действует. Схема определяет параметры его «понимания», влияя на то, как он интерпретирует данные, оценивает варианты и расставляет приоритеты целей [3].

Вследствие этого рациональность ИИ-агента полностью базируется на онтологии, базе данных и логике вопросов и ответов. Это не изолированная конструкция, а диалоговая платформа с функцией обратной связи, в семантике которой содержится диапазон целей, на достижение которых агент ИИ запрограммирован.

Важнейшим аспектом рациональности агента ИИ являются последовательность и согласованность умозаключений, поскольку рациональные решения должны соответствовать последовательной логике, избегая противоречий в процессе рассуждения [4]. Это особенно важно для многоагентных систем. Они классифицируются как рои, работающие для достижения общей цели [5].

Адаптивность также предполагает рациональность, поскольку агентам ИИ необходимо корректировать решения и стратегии по мере появления новой информации, обеспечивая при этом контекстную релевантность и эффективность действий [6]. Этическая основа в виде формализованного аппарата семантики нужна, чтобы агенты ИИ учитывали человеческие ценности, достоинство и права.

В идеале этос агентов ИИ должен сочетать рациональность этической обоснованности и согласованности целей. Эта онтология направляет их автономное поведение, обеспечивая логичность, этическую обоснованность и соответствие решений общественным и системным приоритетам. В данном контексте важная роль отводится атрибутам автономности, сложности принятия решений, адаптивности и потенциалу воздействия. Атрибуты переводят семантику абстрактных принципов рациональности, этической обоснованности и согласованности целей в измеримые факторы, которые могут определять управление и человеческий контроль.

Автономность, которая измеряет степень независимости в принятии и исполнении решений [7], демонстрирует прикладной смысл принципа рациональности, определяя, насколько эффективно агент ИИ может действовать самостоятельно, соблюдая этические ограничения. В медицине [8] способность действовать автономно обеспечивает своевременное вмешательство в условиях ограниченных ресурсов, соблюдая при этом этические ограничения приоритета безопасности и конфиденциальности пациентов, согласует действия с непосредственными клиническими целями.

Чем сложнее процесс принятия решений, тем выше потребность в прозрачности, справедливости и последовательности для предотвращения непредвиденных последствий и обеспечения соответствия решений этическим принципам [9]. Система должна анализировать огромное количество юридических прецедентов, деталей дел и контекстуальных факторов, обеспечивая при этом справедливость и беспристрастность своих рекомендаций [10]. Сложность этой задачи требует прозрачности в принятии решений, четкого объяснения причин рекомендации, а также соблюдения этических принципов недопущения расовой и социально-экономической предвзятости. Неспособность решить проблему сложности принятия решений может привести к несправедливым результатам и подорвать доверие как к агенту ИИ, так и к используемому им юридическому процессу.

Адаптивность означает способность агента ИИ адаптироваться к новым данным и меняющимся обстоятельствам. Это акцент рациональности на отзывчивость и согласованность целей. Атрибут гарантирует, что агент ИИ остается эффективным, контекстуально релевантным и этически последовательным даже в динамичных условиях. Например, агент ИИ, управляющий интеллектуальной энергосистемой [11], должен адаптироваться к меняющемуся спросу на энергию, погодным условиям и использованию возобновляемых источников энергии, уделяя первостепенное внимание эффективности и минимизации воздействия на окружающую среду.

Агент ИИ должен динамически корректировать распределение ресурсов и рекомендовать меры по энергосбережению, не ставя под угрозу критически важные услуги. Такая адаптивность гарантирует эффективность и этическую согласованность принимаемых им решений, согласующихся как с текущими потребностями, так и с долгосрочными целями устойчивого развития. Это концепция согласованности целей, которая оценивает, насколько поведение агента ИИ целеустремленно, и обеспечивает пропорциональный контроль его социальных эффектов. В результате фиксируется способность агента ИИ находить баланс между краткосрочными целями и долгосрочными системными приоритетами.

Заключение

В совокупности атрибуты обеспечивают практическую основу для перевода абстрактных принципов рациональности, этической обоснованности и согласованности целей в измеримые факторы, определяющие управление и контроль со стороны человека. Так, агент ИИ, управляющий городским транспортным потоком, может значительно снизить заторы и выбросы, влияя на миллионы пассажиров и окружающую среду [12]. Он должен учитывать потребности различных заинтересованных сторон, отдавая приори-

тет машинам экстренных служб в часы пик. Это гарантирует, что решения агента ИИ соответствуют краткосрочным целям повышения эффективности дорожного движения и долгосрочным приоритетам повышения общей городской мобильности и качества воздуха.

Список литературы

1. Ethos, Technology, and AI in Contemporary Society: The Character in the Machine / ed.: A. Hess, J. E. Kjeldsen. – NY : Routledge, 2025. – 368 p.

2. Formal verification of ethical choices in autonomous systems / L. Dennis, M. Fisher, M. Slavkovik, M. Webster // Robotics and Autonomous Systems. – 2016. – Vol. 77. – P. 1–14.

3. From Interaction to Impact: Towards Safer AI Agents Through Understanding and Evaluating UI Operation Impacts / Z. J. Zhang, E. Schoop, J. Nichols [et al.] // ArXiv.org. – 2024. – URL: https://arxiv.org/abs/2410.09006 (date of access: 04.09.2025).

4. Elster, J. Rationality / J. Elster // Contemporary Philosophy: A New Survey. – Leiden : Martinus Nijhoff Publishers, 1982. – Vol. 2. – P. 111–131.

5. Multi-modal and multi-agent systems meet rationality: A survey / B. Jiang, Y. Xie, X. Wang [et al.] // ICML 2024 Workshop on LLMs and Cognition. – URL: https://arxiv.org/html/2406.00252v2S7 (date of access: 04.09.2025).

6. MAgIC: Investigation of Large Language Model Powered Multi-Agent in Cognition, Adaptability, Rationality and Collaboration / L. Xu, Z. Hu, D. Zhou [et al.] // Proc. of the 2024 Conf. on Empirical Methods in Natural Language Processing, Miami, Florida, USA, Nov. 2024. – Miami, 2024. – P. 7315–7332. – DOI: 10.18653/v1/2024.emnlp-main.416.

7. Beer, J. M. Toward a framework for levels of robot autonomy in human-robot interaction / J. M. Beer, A. D. Fisk, W. A. Rogers // Journal of Human-Robot Interaction. – 2014. – Vol. 3(2). – P. 74. – DOI: 10.5898/jhri.3.2.beer.

8. Autonomous Artificial Intelligence Agents for Clinical Decision Making in Oncology / D. Ferber, G. O. Wölflein, I. C. Wiest [et al.] // ArXiv.org. – 2024. – URL: https://arxiv.org/abs/2404.04667 (date of access: 04.09.2025).

9. Swanepoel, D. Artificial intelligence and agency: Tie-breaking in AI decision-making / D. Swanepoel, D. Corks // Science and Engineering Ethics. – 2024. – Vol. 30(2). – DOI: 10.1007/s11948-024-00476-2.

10. Uriel, D. Artificial intelligence at the bench: Legal and ethical challenges of informing – or misinforming – judicial decision-making through generative AI / D. Uriel, N. Remolina // Data and Policy. – 2024. – Vol. 6. – DOI: 10.1017/dap.2024.53.

11. Malik, F. H. A review: Agents in smart grids / F. H. Malik, M. Lehtonen // Electric Power Systems Research. – 2015. – Vol. 131. – P. 71–79. – DOI: 10.1016/j.epsr.2015.10.004.

12. Lungu, M. A. Smart Urban Mobility: The Role of AI in Alleviating Traffic Congestion / M. A. Lungu // Proc. of the Intern. Conf. on Business Excellence. – 2024. – Vol. 18, iss. 1. – P. 1441–1452. – DOI: 10.2478/picbe-2024-0118.

УДК 004.8

Пути повышения качества образования с применением технологий искусственного интеллекта

О. Н. Ляховец
Гомельский государственный
колледж кулинарии, Беларусь

Одним из самых актуальных вызовов и одновременно мощных инструментов трансформации в современной образовательной среде является искусственный интеллект. Мы живем в эпоху, когда цифровая революция перестала быть будущим – она стала нашей реальностью. Учащиеся, которые сегодня приходят в аудитории школ, колледжей и вузов, являются цифровыми потребителями. Их восприятие информации, способы коммуникации и обучения кардинально отличаются от тех, что были еще 10 лет назад.

В этих условиях традиционная система образования сталкивается с принципиально новыми требованиями: от педагогов ждут не просто передачи знаний, которые теперь доступны в один клик, а формирования компетенций – критического мышления, креативности, умения работать с большими массивами данных и решать сложные, нестандартные задачи. Здесь возникает главный вопрос, можно ли удовлетворить эти новые запросы, используя лишь старые методики. Ответ, как показывают многочисленные исследования и передовая практика, очевиден – нет.

Технологии искусственного интеллекта – это не просто модный тренд или «цифровая игрушка», это принципиально новый класс инструментов, способных совершить качественный скачок в эффективности и доступности образования. Они позволяют перейти от конвейерной, усредненной модели обучения к гиперперсонализации, где образовательная траектория выстраивается под уникальные потребности, темп и возможности каждого конкретного учащегося.

Также возникает и другой вопрос, может ли робот заменить учителя, вести вместо него уроки. Ответ опять – нет. Искусственный интеллект не может заменить учителя, потому что у него нет эмпатии в человеческом понимании, эмоций и чувств. Ведь учитель должен не только дать знания, но и воспитать духовно-развитую личность. Вместе с тем нейросети – это мощнейший инструмент, который полностью меняет мир или уже его изменил.

Педагогам просто необходимо научиться пользоваться нейросетями, хотя бы для составления плана урока, на который тратится очень много времени. Например, у автора на это в целом уходило 20 мин на один урок.

Все нейросети, которые применялись в работе автором, входят в перечень сетевых сервисов, согласованных с ГИАЦ Министерства образования, утверждены заместителем министра образования Республики Беларусь в методических рекомендациях по использованию технологий искусственного интеллекта в образовательном процессе учреждений общего среднего образования от 02.07.2025 г. Из данного перечня самым лучшим является DEEPSEEK. Многие версии бесплатные, установить их можно легко как на телефон, так и на компьютер.

Итак, вводим запрос: создай подробный план урока английского языка по теме, например, «Супы и технология приготовления» для учащихся профессионально-технического образования на 45 мин с целями и задачами, межпредметными связями. Через 6–8 с виден подробный план урока. Можно попросить создать его и в виде таблицы.

Пример полученного плана урока от нейросети.

План урока английского языка.

Тема: «Soups and Cooking Technology (супы и технология приготовления)».

Целевая аудитория: учащиеся профессионально-технического образования.

Продолжительность: 45мин.

Уровень языка: Pre-Intermediate / Intermediate (A2-B1).

1. Цели и задачи урока.

Общая цель: сформировать и активизировать лексику по теме «Приготовление супов» на английском языке и развить навыки профессионального общения.

Практические задачи:

лексические: ввести, отработать и активизировать новую лексику по теме (названия супов, ингредиентов, стадий приготовления, кухонного инвентаря и глаголов приготовления);

грамматические: активизировать использование повелительного наклонения (Imperative Mood) и пассивного залога (Passive Voice) в контексте рецептов и инструкций;

развивающие:

– аудирование: развить умение извлекать основную информацию из профессиональной видеорецептуры;

– говорение: развить навыки монологической (описание рецепта) и диалогической (работа в парах) речи в профессиональной ситуации;

– чтение: развить умение понимать текст кулинарного рецепта;

– письмо: составить краткий план или рецепт супа.

Воспитательная задача: воспитывать аккуратность, ответственность и умение работать в команде.

2. Межпредметные связи:

«Технология приготовления пищи»: использование знаний о классификации супов, технологии их приготовления, правильной нарезке овощей (techniques of cutting vegetables: dicing, slicing, julienne);

«Товароведение пищевых продуктов»: знание названий и характеристик ингредиентов (овощей, мяса, бульонов, специй);

«Оборудование предприятий общественного питания»: знание названий кухонного инвентаря и посуды (saucepan, stockpot, blender, ladle, chef's knife);

«Санитария и гигиена»: упоминание правил безопасности и гигиены на кухне (food safety, hygiene, to wash hands, to sterilize).

3. Оборудование и материалы:

компьютер/ноутбук, проектор или интерактивная доска;

презентация PowerPoint/Google Slides с ключевой лексикой и изображениями;

раздаточный материал (Handouts) с упражнениями;

видеофрагмент с рецептурой супа на английском (например, классический овощной суп или крем-суп), длительность: 2–3 мин;

аудиозапись для фонетической зарядки.

4. Ход урока (45 мин).

I. Организационный момент (2 мин).

Приветствие: «Good afternoon, everyone! Ready for our cooking class in English?»

Постановка целей: объявление темы и целей урока. «Today we are going to become international chefs! We will learn how to describe the process of making soups in English, learn new professional vocabulary and even watch a video recipe».

II. Этап актуализации знаний (Warm-up) (5 мин).

мозговой штурм: «Let's warm up our brains! What is the first dish that comes to your mind when you hear the word 'SOUP'? Name any soups you know in English or in Russian».

Учащиеся называют супы (борщ, куриный суп, грибной крем-суп и т. д.). Учитель записывает известные варианты на доске (borscht, chicken noodle soup, mushroom cream soup, tomato soup) и добавляет два-три новых (minestrone, gazpacho, French onion soup).

Цель: создать мотивацию и выявить текущий уровень знаний по теме.

III. Введение и отработка новой лексики (15 мин).

1. Визуализация (7 мин):

демонстрируется презентация с тематическими блоками:

types of soups: clear soup, cream soup, broth, chowder, puree soup;

main ingredients: chicken breast, potatoes, carrots, onions, celery, tomatoes, cabbage, beans, noodles;

cooking verbs: to chop, to slice, to dice, to peel, to boil, to simmer, to stir, to add, to season;

kitchenware: saucepan, stockpot, knife, cutting board, ladle, blender.

Каждое слово сопровождается картинкой и транскрипцией.

Преподаватель отрабатывает произношение.

2. Контролируемая практика (8 мин):

раздаточный материал № 1: упражнение «Match the word with the picture/definition».

Пример: соединить глагол to simmer с определением «to cook something just below boiling point» или картинкой, где жидкость слегка пузырится.

Учащиеся выполняют задание индивидуально, затем проверяют в парах. Общая проверка с классом.

IV. Развитие навыков аудирования (8 мин):

1. Pre-Listening (2 мин): «Now we are going to watch a video recipe for a classic vegetable soup. Before watching, look at the list of ingredients. What do you think we will need?» (учащиеся просматривают список ингредиентов на экране или в раздаточном материале).

2. While-Listening (4 мин): учащиеся смотрят видео дважды.

Задание 1 (первый просмотр): проверить, какие ингредиенты из списка были использованы.

Задание 2 (второй просмотр): расставить этапы приготовления (steps of cooking) в правильном порядке (упражнение на установление последовательности).

3. Post-Listening (2 мин), краткое обсуждение: «Was the recipe easy to follow? What was the most interesting step?»

V. Развитие навыков говорения (10 мин).

Ролевая игра «In the Kitchen» (10 мин): учащиеся делятся на пары. Один – шеф-повар (Head Chef), другой – стажер (Trainee Chef).

Задача для шефа: дать инструкции стажеру по приготовлению простого супа (например, картофельного), используя активный словарь и повелительное наклонение. («First, peel and dice the potatoes. Then, bring the water to a boil...»)

Задача для стажера: задавать уточняющие вопросы. («How many potatoes should I peel?», «Should I stir it now?»)

Учитель ходит по классу, оказывает поддержку и корректирует речь. После подготовки одна-две пары представляют свой мини-диалог классу.

VI. Подведение итогов и рефлексия (4 мин).

Обобщение: учитель кратко повторяет ключевые моменты урока: «So, today we have learned the names of different soups, cooking verbs and how to give instructions».

Рефлексия: «What was the most useful for you today? What new word or phrase will you remember?»

Объявление домашнего задания (1 мин):

базовый уровень: написать рецептуру своего любимого супа на английском языке (7–10 предложений), используя новую лексику;

продвинутый уровень: найти рецептуру супа в интернете на английском языке, выписать пять новых слов и пересказать этапы его приготовления.

5. Домашнее задание.

1. Writing Task: write a recipe for your favorite soup. Use at least 10 words from today's lesson (ingredients, verbs, kitchenware, 7–10 sentences).

2. Optional Creative Task: find a picture of an unusual soup online. Be ready to describe it in the next lesson (its color, main ingredients, and how you think it's made).

Итак, все это сильно экономит время педагога. В течение трех недель автор составляла планы уроков с помощью нейросети – 32 учебных занятия, и весь процесс занял 1 ч 36 мин с учетом усложнения части урока, изменения методики преподнесения учебного материала и пр. Самостоятельная подготовка к учебным занятиям заняла в среднем 10 ч 40 мин на это же количество уроков. Разница очевидна.

Здесь важно учесть правильно введенный запрос, а именно план, текст, вопросы и т. д., для какой аудитории, какой объем. Очень удобно с помощью нейросетей составлять тестовые задания. Берем кусок текста, пройденного на уроке, и пишем: сгенерируй, например, 20 вопросов с ответами внизу. Можно запросить несколько вариантов, 50 вопросов или сколько требуется. Из материала, скопированного в нейросеть, через 6–8 с появится сам тест. Конечно, его следует перечитать и перепроверить.

Также можно попросить помочь объяснить сложную тему простыми словами. Например, запрашиваем: объясни простыми словами, когда нужно применять слово poultry в английском языке. Ответ, сгенерированный нейросетью:

Poultry – это домашняя птица, которую разводят для еды.

Проще запомнить так:

Poultry = «птица с тарелки».

Когда его нужно применять?

Используйте слово poultry, когда говорите о птицах как о продукте питания:

курица (chicken),
индейка (turkey),
утка (duck),
гусь (goose),
перепел (quail).

Примеры в предложениях:

1. В магазине:

«I am going to the supermarket to buy some poultry» (я иду в супермаркет купить птицы). Имеется в виду курица, индейка и т. д.

«The poultry section is next to the fish counter» (отдел с птицей рядом с рыбным прилавком).

2. В диете или готовке:

«We eat poultry twice a week» (мы едим птицу два раза в неделю).

«Poultry should be cooked thoroughly» (птицу нужно готовить до полной готовности).

3. В сельском хозяйстве:

«My grandfather's farm raises poultry» (на ферме моего дедушки разводят домашнюю птицу).

Самая главная хитрость:

poultry – это мясо, а не животное в поле.

Если вы видите живую курицу, гуляющую по двору, это a chicken (курица).

Если вы видите кусок мяса этой курицы на вашей тарелке, можно обобщенно назвать его poultry.

Краткая шпаргалка:

Если говорите о...	То используйте слово...
мясе птицы вообще (курица, индейка, утка):	poultry
конкретно о курице (живой или как о мясе):	chicken
конкретно об индейке:	turkey
конкретно об утке:	duck

Нейросеть сканирует открытые базы данных, научные статьи и образовательные платформы, подбирая релевантные материалы. При этом экономится время на ручной поиск. Она создает индивидуальные задания, учитывая уровень знаний, интересы и темп обучения каждого учащегося, например, для слабых учащихся – дополнительные задачи, для продвинутых – проектные задания.

Также можно проанализировать успеваемость учащихся с помощью DEEPSEEK, скопировав не просто ведомость успеваемости за полугодие,

а сам вордовский документ. За менее 10 с нейросеть сделает подробный анализ успеваемости каждого учащегося. Однако не стоит загружать файл с фамилиями в нейросеть, их следует изменить на номера по списку. Необходимо соблюдать конфиденциальность личных данных и не распространять их в нейросети.

В своей работе автор применяла и платформу QUIZIZZ, которая также была создана с помощью искусственного интеллекта. Эта платформа помогает учащимся на учебном занятии в виде теста или викторины отвечать на вопросы. В процессе работы с тестом педагог в своем компьютере видит, кто какие ошибки допустил, сколько процентов знаний усвоил, какую отметку получил и с чем надо работать дальше. Тесты, сгенерированные нейросетью, можно распечатать, и учащиеся будут отвечать на вопросы уже письменно.

На создание тестового задания самостоятельно уходит примерно 10 мин, с помощью нейросети – меньше минуты. Создав 16 тестовых заданий за три недели, автор потратила около 20 мин, без использования нейросетей на это уходило около 2 ч 40 мин.

Также можно использовать нейросети для генерации изображений. Для учебного занятия можно легко сгенерировать картинку на заставку. Например, дать задание создать картинку – карту меню из морепродуктов на английском языке, и нейросеть создаст эту картинку по вашему запросу.

С помощью нейросети KANDINSKY на уроках английского языка можно предложить учащимся описать изображения, например, по теме «Молочные продукты». Учащиеся на компьютере придумывают описание картинки, используя лексику урока, генерируют изображение с помощью нейросети, затем в паре устно описывают картинку друг другу.

Еще одной идеей был визуальный словарь. Учащиеся создают изображения по словам, затем стараются угадать слово. Таким образом, пока учащиеся пишут запрос в нейросети, они тренируют грамматику и запоминают слова, улучшают навыки письма.

Итак, можно выделить главные положительные стороны работы с искусственным интеллектом:

– экономится время, подходит каждому учащемуся;

– быстрая обратная связь, т. е. учащиеся сразу видят ошибки в тестах или упражнениях и им не нужно ждать, пока учитель все проверит;

– интересные учебные занятия;

– искусственный интеллект следит за процессом, т. е. автоматически анализирует, кто отстает, а кто преуспевает, и учитель заранее видит проблемы и может помочь;

– учащиеся могут учиться дополнительно дома в чат-ботах, на платформах;

– уменьшается бумажная работа, больше времени отводится процессу преподавания.

Недостатки:

– если искусственный интеллект заменяет живое общение, например, чат-бот вместо учителя, то с учащимися можно потерять эмоциональную связь;

– иногда ИИ ошибается, дает некорректные ответы и учителю приходится их перепроверять;

– не все имеют хороший интернет и компьютеры на каждом учебном занятии;

– искусственный интеллект искореняет творчество, например сочинения проверяет «по шаблону», а не оценивает их оригинальность.

Однако время не стоит на месте, мир развивается. Педагоги так или иначе будут вынуждены подстраиваться под современные реалии жизни и с помощью искусственного интеллекта учиться упрощать поставленные задачи и получать преимущества для дальнейшей работы. Лучше начинать с малого – внедрять искусственный интеллект в свою работу поэтапно, осознанно, дозированно. Лучше всего – сочетать технологии с человеческим подходом.

Как говорил Сократ: «Образование – это зажигание пламени, а не наполнение сосуда».

УДК 004.853

Обучение учителей математики использованию технологии искусственного интеллекта для разработки учебно-методических материалов

Д. И. Прохоров✉
Минский городской институт
развития образования, Беларусь
E-mail: prohorov@minsk.edu.by

Введение

Инновационный многовекторный процесс цифровой трансформации системы дополнительного педагогического образования ставит перед научным сообществом вопросы, связанные с разработкой, внедрением и методическим сопровождением использования интернет-технологий для повышения квалификации и активизации самообразовательной деятельности учителей математики. В последние годы искусственный интеллект (ИИ) все больше интегрируется в нашу повседневную жизнь, предоставляя социуму разнообразные возможности, улучшая качество жизни и способствуя самообразованию и самореализации.

Автором выявлено противоречие между запросом общества в непрерывном совершенствовании профессиональных компетенций, обогащении педагогического и методического опыта в условиях цифровизации и недостаточной разработанностью технологий и методик целостного концептуально обоснованного повышения квалификации и активизации самообразовательной деятельности (далее – ПКиСД) включая межкурсовой период с использованием технологий ИИ.

Методологическое основание использования ИИ в процессе ПКиСД

Научным базисом проектирования и разработки ПКиСД учителей математики с использованием ИИ выступают взаимнообусловленность и взаимодополнение *полипарадигмального подхода*:

– синергетического: учет динамики и механизмов синергетических оснований процесса ПКиСД учителей математики (гомеостатичность, иерархичность, нелинейность, открытость, вариабельность, эмерджентность, коммуникативность); опора на формы, методы, средства и ресурсы поиска нового педагогического и математического знания, которое имеет про-

блемно-эвристический характер и базируется на интуиции, воображении и творчестве учителей математики;

– системно-деятельностного: организация активной, проблемной, эвристической групповой и самостоятельной познавательной деятельности; учет особенностей организации тьюторства и инженерии знаний слушателей для организации конструктивной деятельности в процессе ПКиСД учителей математики;

– праксеологического: оптимизация процесса обучения учителей математики, трансляция опыта на их профессиональную деятельность; выявление условий наиболее эффективного повышения уровня их профессиональных компетенций, причин, позволивших или не позволивших достичь запланированной цели обучения;

– компетентностного: непрерывное усиление, обогащение и углубление профессиональных компетенций учителей математики; практико-ориентированный характер обучения, увеличение роли самостоятельной деятельности слушателей по решению педагогических задач;

– логистического: управление логистическими потоки процесса ПКиСД учителей математики (внешний контур – поток трудовых ресурсов, финансовый и материально-технический потоки; внутренний контур – информационный и учебно-методический потоки); снижение субъективных рисков организации и проведения повышения квалификации;

– коннективистского: использование активных и интерактивных форм, методов и средств обучения учителей математики; создание условий для профессиональной коммуникации слушателей и преподавателей с использованием веб-ориентированных ресурсов; подбор и разработка специальных веб-ориентированных ресурсов обучения с учетом требований дидактического дизайна;

– инструментального: расширение инструментария учителей математики по структурированию, обобщению учебной информации посредством дидактических многомерных инструментов и интернет-технологий (инфографика, логико-смысловые модели и т. д.); интериоризация полученных в процессе ПКиСД теоретических и практических знаний.

Что является основанием для определения *общедидактических принципов обучения*:

– фундаментальности, которая определяет цель и задачи: сохранение информационной ценности в течение всей профессиональной деятельности педагога; отражение современных тенденций развития методики преподавания математики; учет инноваций в области ИИ;

– гуманизации и гуманитаризации, которые определяют формы, методы обучения, коммуникацию между преподавателем и слушателями с использованием веб-ориентированных ресурсов: опора на мотивационно-

ценностную составляющую процесса обучения; структурирование содержания учебных программ повышения квалификации; организация форм, методов и средств обучения; субъект-субъектное взаимодействие между преподавателем и слушателями;

– опережающего характера обучения, который организует содержание, формы и методы обучения и коммуникации: построение проблемной структуры учебной информации в веб-ориентированных ресурсах обучения при ведущей роли теоретического знания и практических навыков; обучение строится на высоком, но доступном для слушателей уровне сложности, темпе; дифференцированность управления самообразовательной деятельности учителя математики; диалоговые, полилоговые и фасилитационные формы взаимодействия.



Рис. 1. Структурно-логическая схема освоения содержания обучения при ПКиСД учителей математики

Выделение *принципов, отражающих специфику системы дополнительного образования взрослых*, также носящих общедидактический характер:

– партисипативности, устанавливает требования к коммуникации между преподавателем и слушателями: учет мнения каждого слушателя на основе эмпатии; проведение систематических консультаций; учет личност-

ного и профессионального опыта учителей математики; делегирование определенных управленческих полномочий слушателям; непрерывная обратная связь;

– фундирования, определяет требования к ресурсам обучения: представление содержания в виде «спиралевидной конструкции» [1] с четырьмя витками с нарастанием сложности без увеличения объема работы; каждый виток разворачивается в три пласта фундирования (содержательный, информационно-технический, методический). Структурно-логическая схема освоения содержания обучения при ПКиСД учителей математики представлена на рис. 1;

– оптимальной информационной насыщенности, определяет содержание обучения, формы, методы и средства обучения: послойное распределение учебной информации с нарастающей сложностью заданий; использование веб-ориентированных ресурсов; организация проблемного и эвристического обучения.

Электронное средство обучения «Дидактический дизайн преподавания математики в учреждениях общего среднего образования»

С целью обеспечения учебно-методического сопровождения процесса ПКиСД учителей математики по освоению современных методик и технологий обучения, в том числе касающихся внедрения технологий ИИ в образовательный процесс учреждений общего среднего образования, был разработан ресурс [2], который включает (рис. 2):

– глоссарий (более 180 понятий и терминов);

– формы установления взаимодействия, учебной и профессиональной коммуникации преподавателя со слушателями (учителями математики): e-mail преподавателей, форум «Вопрос-Ответ» (обсуждение заданий, методическая поддержка, техническая поддержка), чат-бот «Профессор МГИРО» (@professorMGIRO_bot) – телеграм-чат на основе ИИ;

– материалы для освоения содержания программы ПКиСД учителями математики «Дидактический дизайн преподавания математики в учреждениях общего среднего образования»: вопросы государственной политики в сфере цифровизации образования (две темы), научно-теоретические основания структурирования содержания обучения математике (пять тем), основы работы со специализированным программным обеспечением (11 тем);

– примеры фрагментов дидактических сценариев учебных занятий, отражающих методику использования дидактических многомерных инструментов при обучении математике (база постоянно пополняется материалами, разработанными учителями математики в процессе ПКиСД);

– коллекцию готовых учебных математических апплетов и методические указания по их использованию на учебных занятиях по математике на II–III ступенях общего среднего образования;

– фрагменты видеолекций и практических занятий, где даются инструкции по работе с компьютерными программными средствами: редакторами векторной графики, динамическими программными средами, конструкторами инфографики, учебными математическими апплетами, GPT-чатом и т. д., а также алгоритмы разработки собственных веб-ориентированных ресурсов;

– материалы для текущей и итоговой аттестации (дифференцированные по трем уровням сложности).



Рис. 2. Структурно-логическая схема веб-ориентированного ресурса обучения

Каждая учебная тема, размещенная на авторском ресурсе, состоит из основного материала, дополнительной литературы, заданий для самостоятельного выполнения. Содержательный блок авторского ресурса, обеспечивающего повышение ПКиСД учителей математики, имеет спиралевидную конструкцию из четырех витков:

1. Учителя математики под руководством преподавателя осваивают алгоритмы структурирования учебной информации для учащихся II–III ступеней общего среднего образования с использованием возможностей дидактических многомерных инструментов (таблиц, блок-схем и т. д.) и простых веб-ориентированных ресурсов (веб-презентаций, онлайн интерактивных досок и др.). Например, «Используя возможности Google Disk, создайте электронную таблицу, “Свойства тригонометрической функции $f(x) = \sin x$” (область определения/значения, четность/нечетность, периодичность, промежутки возрастания/убывания, достижение максимального/минимального значения, интервалы положительных/отрицательных значений, нули функций)».

2. Учебная информация предыдущего витка обогащена алгоритмами решения педагогических задач на основе использования различных существующих веб-ориентированных ресурсов, применения на учебных занятиях по математике информационно емких визуальных изображений (учебных математических апплетов, инфографики, структурно-логических схем и т. д.) по отдельным учебным темам в готовом виде. Например, «Используя созданную ранее электронную таблицу и возможности онлайн-редактора https://canva.com, создайте инфографику по теме “Свойства тригонометрической функции $f(x) = \sin x$”. Дополните созданную инфографику ссылками на учебные математические апплеты библиотеки сайта https://urok.1c.ru.».

3. Содержание витка дает возможность учителям математики под руководством преподавателя самостоятельно разрабатывать веб-ориентированные ресурсы, рассматривать различные методические аспекты обучения учащихся навыкам обобщения и структурирования учебной информации по укрупненным темам (логико-смысловые модели, учебные математические апплеты и т. д.). Например, «Используя возможности динамического программного онлайн-сервиса 1С: “Математический конструктор”, создайте учебный математический апплет “Тригонометрическая функция (синус)”, в котором отображается график тригонометрической функции вида $f(x) = a\cdot\sin(bx + c) + d$, где a, b, d – параметры со значениями от –5 до 5, шаг изменения 0,1, c – параметр со значением от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$, шаг изменения $\frac{\pi}{12}$. Настройте апплет так, чтобы в записи тригонометрической функции $f(x) = a\cdot\sin(bx + c) + d$ отображались значения параметров (коэффициентов) a, b, c, d. Например, $f(x) = -\sin\left(2\cdot x + \frac{\pi}{3}\right) + 1$».

4. Предназначен преимущественно для самостоятельного освоения учителями математики способов построения процесса обучения математике с использованием веб-ориентированных ресурсов, что позволяет слушателям самостоятельно обобщить и транслировать свой педагогический опыт коллегам. Например, «Дополните учебный математический апплет $f(x) = a\cdot\sin(bx + c) + d$ еще одним параметром вида $x = e$, где e – параметр со значениями от –5 до 5, шаг изменения 0,1. Оформите апплет так, чтобы с его помощью можно было визуализировать решение тригонометрического уравнения вида $a\cdot\sin(bx + c) + d = e$, где a, b, d, e – действительные числа от –5 до 5, шаг изменения 0,1, c – параметр со значением от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$, шаг

изменения $\frac{\pi}{12}$. Разработайте сценарий проведения учебного занятия с использованием данных апплетов. Перешлите коллегам для экспертизы разработанный вами апплет и сценарий учебного занятия».

По мнению автора, в процессе ПКиСД учителей математики целесообразно пользоваться теми методами и средствами ИИ, которые будут полезны в профессиональной деятельности, а также являются общезначимыми и могут использоваться в обыденной жизни (например, чат-боты). Телеграм-бот «Профессор МГИРО» (@professorMGIRO_bot) разработан на основе технологии нейросетей Yandex, свободно ведет диалог по содержанию учебной программы повышения квалификации учителей математики по теме «Дидактический дизайн преподавания математики в учреждениях общего среднего образования». Кроме общедоступных баз методической литературы и цифровых образовательных технологий, в телеграм-бот добавлены дополнительные векторные базы: 12 пособий по методике преподавания математики и внешние базы данных, касающиеся применения информационно-коммуникационных технологий в образовательном процессе.

Пример заданий по разработке учебно-методических материалов при помощи технологий ИИ

При работе с сервисами генеративного ИИ главной задачей пользователя является корректная формулировка задания – промптинг (англ. prompting). Рассмотрим возможности ресурсов ИИ по созданию учебно-методического материала на примере занятия по теме «Геометрические иллюзии», V класс. Как было отмечено ранее, основная работа пользователя заключается в формулировании наиболее подробного и понятного для нейронной сети запроса.

Нейронные сети используют разнообразные источники информации, включая открытые сетевые ресурсы. Прежде чем применять сгенерированные ими материалы в образовательном процессе, их необходимо тщательно проверить и критически проанализировать, оценив соответствие запросу, достоверность, актуальность, объективность и полноту. Прежде всего, требуется определить цели проведения данного мероприятия. Цели могут быть сформулированы следующим образом [3]. По завершении занятия учащийся будет знать определение понятия «геометрические иллюзии»; понимать причины возникновения геометрических иллюзий у человека; применять алгоритм создания простых геометрических иллюзий при помощи листа бумаги и карандашей; анализировать причины возникновения геометрических иллюзий в архитектуре и художественных произведениях; синтезировать отдельные элементы изображений, которые в совокупности

приводит к появлению геометрических иллюзий, объяснять их причины; критически оценивать изображения на предмет их правдоподобия (на примере невозможного треугольника Пенроуза, куба Эшера и т. д.).

Разработаем сценарий занятия на основе сервиса Deepseek (https://deepseek.com). Перейдите на страницу инструмента ИИ, установите русскоязычную рабочую область, выберите раздел «Начать разговор». При помощи кнопки «Добавить» текстовый файл можно загрузить несколько разработанных ранее сценариев занятий для того, чтобы обучить нейронную сеть привычному стилю написания и оформления учебно-методических материалов.

В поле ввода запроса введите следующий текст: Разработай сценарий занятия для 5-го класса по теме «Геометрические иллюзии» на 45 минут. На организационном этапе добавь интерактивный метод стимулирования мотивации учения и активизации когнитивных процессов учащихся. На этапе изучения нового материала дай определение понятия «геометрическая иллюзия», приведи примеры геометрических иллюзий в окружающем мире. Опиши с точки зрения физиологии и психологии почему у человека возникают геометрические иллюзии. Напиши про геометрические иллюзии в архитектуре Древнего Египта и Древней Греции. Дай краткое описание биографии художников-иллюзионистов. На этапе обобщения и закрепления изученного материала напиши алгоритм создания простых геометрических иллюзий при помощи листа бумаги и карандашей. Обобщи предыдущий материал в краткую памятку. На этапе подведения итогов занятия добавь интерактивный метод учебной рефлексии.

Скопируйте полученный текст, по вашему усмотрению проведите редактирование, коррекцию и форматирование материала.

Добавим диагностический инструментарий для контроля усвоения материала в конце учебного занятия. Перейдите на страницу инструмента Diffit (https://app.diffit.me). В разделе «Введите тему, тему или вопрос» поместите теоретические сведения, знание учащихся которых вы хотите проверить в конце занятия. Установите настройки: Уровень чтения – 5 класс, Язык – Русский, Тип чтения – Информационный текст. Сгенерируйте тестовые задания, проверьте их корректность, соответствие материалам сценария и верность формулировок самих вопросов и вариантов ответов на них. В разделе «Вопросы с краткими ответами» сформулируйте практические задания, предназначенные для самостоятельного выполнения учащимися. При необходимости представьте диагностический инструментарий в виде раздаточных карточек (раздел «Карта пузырьков» с рабочей тетрадью изображений в правой части экрана), скачайте созданные объекты. Дополните разработанный ранее сценарий занятия диагностическим инструментарием, с этим учетом скорректируйте содержание его этапов.

Соберем визуальные материалы: при помощи интернет-поисковиков найдите и скачайте картинки по запросам «геометрические иллюзии на плоскости», «геометрические иллюзии в пространстве», «геометрические иллюзии в архитектуре», «иллюзия Мюллера-Лайера», «картины Оскара Реутерсварда», «лестница Эшера», «комната Эймса в миниатюре», «невозможный треугольник Пенроуза», «куб Эшера», «куб Неккера», «ваза Рубина» и т. д.

Дополним ресурсное обеспечение учебного занятия дидактическими многомерными инструментами: на основе ресурса Leonardo.AI (https://app.leonardo.ai) создайте инфографику, для чего в строку запроса введите понятия, аналогичные тем, которые задавались для интернет-поисковиков. Выберите соответствующие рисунки и произведите необходимые настройки в левой части экрана, скачайте сгенерированные изображения. При помощи интернет-ресурса LearningApps.org разработайте несколько интерактивных учебных приложений по теме «Геометрические иллюзии». Включите в сценарий занятия описания технологии использования разработанных дидактических многомерных инструментов на соответствующих этапах его проведения.

Объединим визуальные материалы и дидактические многомерные инструменты в презентацию при помощи сервиса Gamma AI (https://gamma.app). На главной странице ресурса выберите раздел «Вставить текст». В появившемся окне ввода запишите запрос, аналогичный описанию сценария занятия, который формулировали для сервиса Deepseek. Выберите пункт Презентация, Стиль страницы – По умолчанию, нажмите на кнопку «Продолжить». В левой части экрана выберите необходимые настройки: Текстовое содержание – Сгенерировать, Количество текста на одной карте – Краткое описание, нажмите на кнопку «Продолжить». В правой части экрана выберите Тип темы – Светлая и Профессия, Тема – Tranquil, нажмите на кнопку «Сгенерировать». Обратите внимание, что содержание и оформление презентации можно изменять при помощи панели инструментов в правой части экрана.

Поскольку сервисы Deepseek и Gamma AI используют разные нейронные сети, то идентичные запросы для создания сценария занятия и соответствующей презентации могут давать различный результат. Более того, один и тот же запрос в один и тот же сервис ИИ может приводить к разным результатам. Причины таких различий кроются в особенностях функционирования нейронных сетей.

Приведите в соответствие содержание презентации и разработанного ранее сценария занятия по теме «Геометрические иллюзии» в V классе. Дополните презентацию QR-кодами для прямого доступа к дидактическим многомерным инструментам. Просмотреть итоговый вариант презентации можно нажав на кнопку «Присутствие» в верхнем меню. Нажмите на

кнопку «Поделиться» и выберите команду «Экспорт» для того, чтобы скачать созданную презентацию. Самостоятельно изучите возможности других сервисов ИИ, представленных в параграфе.

Использование инструментов ИИ влечет за собой существенные риски для учащихся, включая воздействие нежелательного материала, а также возможность манипулирования. Кроме того, системы генеративного ИИ, имитирующие взаимодействие с людьми, могут оказывать нежелательное психологическое воздействие на учащихся, что вызывает беспокойство относительно их когнитивного развития и эмоционального благополучия.

Заключение

Проведенный анализ подтверждает, что цифровая трансформация дополнительного педагогического образования требует разработки целостной концепции повышения квалификации и активизации самообразовательной деятельности учителей математики. В качестве методологической основы такого процесса успешно выступает комплекс взаимодополняющих подходов (синергетический, системно-деятельностный, праксеологический, компетентностный, логистический, коннективистский, инструментальный), что обеспечивает его системность, практическую ориентированность и учет специфики взрослой аудитории. Ключевым методическим элементом обучения учителей математики является разработанное электронное средство обучения, которое предоставляет учителям структурированные материалы, инструменты для профессиональной коммуникации и уникальный чат-бот на базе ИИ для методической поддержки. Предложенная спиралевидная модель освоения содержания обучения учителей математики, состоящая из четырех витков, позволяет поэтапно наращивать сложность заданий – от простого структурирования информации до самостоятельной разработки цифровых учебных ресурсов и обмена педагогическим опытом.

Список использованных источников

1. Бровка, Н. В. Интеграция теории и практики обучения математике как средство повышения качества подготовки студентов / Н. В. Бровка. – Минск : БГУ, 2009. – 242 с.

2. Дидактический дизайн преподавания математики в учреждениях общего среднего образования: методическое обеспечение учебной программы повышения квалификации / разраб.: Д. И. Прохоров // Минская городская платформа дистанционного обучения / Мин. гор. ин-т развития образования. – URL: https://do.minsk.edu.by/course/view.php?id=8832#section-12 (дата обращения: 24.09.2025).

3. Мурзагалиева, А. Е. Учебные цели согласно таксономии Блума : сб. заданий и упражнений / А. Е. Мурзагалиева, Б. М. Утегенова. – Астана : Центр пед. мастерства Назарбаев Интеллектуал. шк., 2015. – 54 с.

УДК 004.8:37.0

Перформативные аспекты внедрения искусственного интеллекта в образовательное пространство Беларуси

В. Н. Пунчик
Республиканский институт высшей школы,
Минск, Беларусь

К. В. Дорожкин
Центр искусственного интеллекта и науки о данных,
Санкт-Петербург, Россия

Введение

Искусственный интеллект (ИИ) в последние годы перестал быть прерогативой узкого круга исследователей и стремительно превратился в повседневный инструмент, проникающий во все сферы общественной жизни – от экономики и медицины до образования и культуры. Логика жизненного цикла инноваций позволяет констатировать, что сегодня государства и корпорации конкурируют не столько за создание новых архитектур ИИ, сколько за способы их эффективной интеграции в существующие социальные и образовательные институты.

Для Республики Беларусь данный процесс имеет двойственную природу. С одной стороны, академические традиции концентрировались преимущественно на фундаментальных теоретических аспектах ИИ: математических моделях, алгоритмах машинного обучения, развитии национальных научных школ ИИ и их «закрытом» применении, замкнутом на определенные производственные процессы. С другой – реальность, детерминированная массовым доступом к LLM и нейросервисам, уже диктует новые правила: педагоги используют генеративные сервисы для подготовки уроков, студенты применяют нейросети для написания проектов, бизнес автоматизирует коммуникации с клиентами с помощью чат-ботов. Возникает противоречие: ИИ-инновации распространяются стихийно, а институциональное регулирование остается фрагментированным при наличии фундаментального национального базиса.

В данном контексте особую значимость приобретает актуализация требований к социальному заказу на кадры в области ИИ. Становится очевидным, что речь идет не только о подготовке программистов и аналитиков данных, но и о формировании у каждого специалиста ключевых компетенций XXI века – умения эффективно взаимодействовать с нейросетя-

ми, интерпретировать результаты их работы, сохранять методологическую строгость и этическую ответственность.

Цель настоящей статьи – рассмотреть перформативные аспекты внедрения ИИ в образовательное пространство Беларуси, сопоставив академическую повестку и реальную практику, выявив барьеры и предложив пути институционализации инноваций на основе опыта Российской Федерации. В статье предпринимается попытка соединить академическую рефлексию и прикладные практики, показать, каким образом ИИ становится не только объектом изучения, но и инструментом формирования новой образовательной парадигмы в Беларуси.

На протяжении долгого времени ИИ, появившийся как концепция еще в 1955 г., в научно-педагогическом дискурсе рассматривался преимущественно как абстрактный феномен, объект фундаментальных исследований, сосредоточенных на его математических основах, этических импликациях и философских интерпретациях. Академические форумы концентрировались в основном на построении теоретических моделей и гипотез, что закономерно приводило к разрыву между научным осмыслением ИИ и его стремительным проникновением в образовательную и социально-экономическую практику. Такой подход, в рамках которого ИИ воспринимался как достояние узкого круга IT-специалистов, не учитывал его эволюцию в универсальный инструмент, доступный широким профессиональным и социальным группам.

Современный этап цифровизации образования и науки характеризуется уже не столько внедрением отдельных ИИ-технологий, сколько очередным культурадигмальным сдвигом. Как отмечал автор культурно-праксеологической концепции И. И. Цыркун [1], процесс распространения инноваций имеет две ключевые составляющие:

– естественную, когда освоение технологий происходит стихийно, под влиянием рынка, общества и локальных инициатив;

– искусственную, предполагающую целенаправленное управление со стороны государства, научного сообщества и профильных организаций.

Сегодня в Беларуси наблюдается активное развитие первой составляющей: педагоги, учащиеся, бизнес-структуры и исследователи самостоятельно осваивают возможности генеративных моделей, интегрируя их в учебный и профессиональный процесс. Однако именно искусственная составляющая, связанная с институциональной интеграцией, остается в стадии догоняющего развития.

Обзор тематики крупнейшего Международного форума IT-Академграда «Искусственный интеллект в Беларуси» (2022–2024 гг.) [2–4] позволяет зафиксировать четкую тенденцию дифференциации: если ранее в центре внимания находились информационные технологии в целом, то сегодня

ключевые секции посвящены машинному обучению, нейронным сетям, генеративному ИИ, обработке естественного языка и компьютерному зрению. При этом значительная часть докладов связана с применением ИИ в медицине, промышленности, городской инфраструктуре и образовании. Особый интерес вызывает проект belai.by, который претендует на роль центра консолидации белорусского ИИ-сообщества, объединяя исследователей, разработчиков и практиков. Однако при всей важности горизонтального взаимодействия и объединения такая консолидация сообщества лишь подчеркивает острый закономерный вопрос, кто и каким образом должен готовить специалистов, способных эффективно работать с ИИ для пополнения этого растущего сообщества.

Анализ материалов форумов и сложившейся практики показывает, что ключевые кадровые вызовы можно сгруппировать следующим образом [5]:

– технологические: быстрый темп развития ИИ-технологий приводит к постоянному устареванию учебных программ, а также требует интеграции в обучение облачных решений, инструментов анализа больших данных, генеративных моделей;

– концептуально-дидактические: наблюдается нехватка преподавателей с актуальными компетенциями, слабая трансдисциплинарная координация, а также дисбаланс между теорией и практикой;

– отраслевые: сохраняется разрыв между академическим образованием и реальными потребностями отрасли; недостаточно внимания уделяется развитию soft skills, включая этику и навыки промпт-инжиниринга.

Таким образом, спрос на специалистов с актуальными ИИ-компетенциями растет быстрее, чем система образования успевает обновлять их содержание. Активно внедряемая в стране концепция «Университет 3.0» (предпринимательского университета) демонстрирует свою ограниченность в этом контексте, поскольку ее фокус на коммерциализации инноваций не охватывает всю сложность современных кадровых вызовов. Для подготовки кадров, адекватных вызовам ИИ-эпохи, университетам необходимо, осуществляя цифровизацию Университета 4.0, уже сейчас ориентироваться на принципы Университета 5.0, которые означают не простое наращивание технологий, а их синтез с гуманитарным знанием с акцентом на человекомерность, человекоцентричность и этику ИИ. При этом каждый вуз, разрабатывая свою стратегию развития, должен не только учитывать специфику отрасли и уникальные запросы рынка труда, но и обеспечить студентам возможность освоения лучших практических наработок, созданных в ИИ-сфере этой отрасли. В рамках данной трансформации партнерство с индустрией становится основным аспектом не только развития, но и функционирования вуза. Реализация принципов Университета 4.0 предполагает активное внедрение таких технологий, как Big Data

и ИИ, для повышения общей эффективности и капитализации отраслевых знаний. Одновременно с этим кардинально меняется модель обучения: происходит переход к студентоцентрированному подходу, основанному на проектной работе, которая интегрируется в реальные задачи отрасли, обеспечивая студентам возможность освоения лучших практических наработок.

Между тем практика освоения ИИ-технологий на местах идет гораздо быстрее и носит преимущественно стихийный характер. В школах и вузах уже массово применяются генеративные сервисы (YandexGPT, GigaChat, DeepSeek и др.) для поиска информации, автоматизации рутинных задач, а также для подготовки учебных материалов, презентаций и видеоконтента. Этот процесс, обусловленный личной инициативой педагогов и учащихся, формирует уникальную ситуацию, когда повседневная практика опережает официальную теорию и методологию. Таким образом, значительный объем актуальных компетенций формируется за пределами структурированных академических программ, переходя в область самостоятельного и неформального обучения, при этом не получая должной фиксации, систематизации и накопления ни в рамках образовательной системы, ни в рамках отрасли.

Безусловно, национальная система образования оперативно реагирует на вызовы времени: Министерство образования Республики Беларусь предприняло значимые шаги, включая официальное разрешение на использование нейросетей и внедрение специального курса «Основы использования технологии искусственного интеллекта» для студентов всех специальностей УВО. Кроме того, в стране уже действует масса талантливых инициативных педагогов и внешних IT-школ, которые активно ведут курсы по ИИ. Однако, чтобы этот ценный практический опыт перестал быть стихийным и получил должную фиксацию и капитализацию, его необходимо тщательно изучить и систематизировать. Работа по сбору и валидации фактических наработок должна вестись не только в академической среде, но и непосредственно в сотрудничестве с отраслью, что является критически важным условием для обеспечения актуальности формируемых компетенций.

Такой подход к систематизации наработок критически важен, поскольку позволяет преодолеть дисбаланс в содержании образования. В то время как традиционные образовательные программы основное внимание уделяли прежде всего пониманию теоретических основ нейросетей (как они устроены), в реальной жизни востребованы совершенно иные, инструментальные компетенции, которые формируются стихийно, вне академической среды. К ним относятся умение формулировать эффективные промпты, проверять достоверность и релевантность ответов, корректно работать с «галлюцинациями» моделей, а также соблюдать этические и пра-

вовые нормы. Эти навыки ближе к освоению офисных приложений, чем к фундаментальной науке, и их фактическое отсутствие в формальном обучении ставит перед системой стратегический вопрос: *как включить новые, критически важные компетенции в программы подготовки специалистов?* Оставленный без внимания, этот стихийный процесс приведет к углублению разрыва между уровнем подготовки выпускников и требованиями рынка. Решение видится в разработке новых образовательных программ и методик, которые будут ориентированы не на глубокое изучение архитектур ИИ, а на практическое владение инструментами генеративного ИИ с учетом специфики решаемых профессиональных задач, развитие критического мышления при работе с результатами моделей и формирование устойчивых этических норм взаимодействия с ИИ.

Особое значение приобретает трансдисциплинарность: разработка таких программ должна вестись в тесном взаимодействии практиков из отрасли, педагогов, IT-специалистов, юристов и психологов. Только в этом случае можно будет подготовить специалистов, обладающих не только техническими, но и социально-гуманитарными компетенциями, необходимыми для ответственного использования искусственного интеллекта. Отсюда вытекает очевидная необходимость создания межведомственных консорциумов, поскольку никакое отдельное учреждение не способно в одиночку обеспечить столь комплексную подготовку.

1. Консорциум по профилям подготовки (трансдисциплинарный фокус). Такой формат нацелен на обеспечение трансдисциплинарности и объединяет вузы и организации в зависимости от типа формируемых компетенций:

технические направления: (Data Science, ML, инженерия ИИ) – для обеспечения фундаментальной и прикладной ИИ-базы;

гуманитарные направления: (этика ИИ, цифровая педагогика, правовое регулирование) – для интеграции социально-гуманитарных компетенций;

прикладные (отраслевые) направления – отдельные специализированные кластеры (например, «ИИ в медицине», «ИИ в финансах», «ИИ в образовании» и «ИИ в промышленности») – для отработки практических кейсов и освоения отраслевых наработок в каждой конкретной сфере.

2. Кластерно-отраслевой консорциум (по научным и предметным областям):

– университеты и НИИ объединяются в профильные кластеры: «ИИ в медицине», «ИИ в образовании», «ИИ в промышленности», «ИИ в культуре и медиа»;

– каждый кластер отвечает за подготовку специалистов, разработку учебных программ и проведение исследований в своей сфере.

Оба формата предполагают участие не только образовательных учреждений, но и бизнеса, государственных органов и профессиональных со-

обществ. Такой подход позволит связать фундаментальную подготовку с реальными запросами индустрии.

Материалы форумов последних лет ясно показывают: вопрос подготовки кадров звучит все громче, но пока остается на периферии основной дискуссии, уступая место обсуждению технологических кейсов. Между тем именно кадровая проблема может стать узким местом, ограничивающим темпы внедрения ИИ в Беларуси. Поэтому необходим переход от обсуждений к институциональным решениям:

– создание национальной программы подготовки кадров в сфере ИИ;

– формирование консорциумов вузов, НИИ и компаний;

– разработка новых образовательных стандартов, включающих навыки промпт-инжиниринга, цифровой этики и междисциплинарных soft skills;

– интеграция белорусских инициатив в международные образовательные проекты.

Безусловно, национальная экосистема ИИ в Беларуси формируется. Один из наиболее значимых сдвигов последних лет связан с изменением статуса ИИ в образовательной среде. Если еще недавно ИИ рассматривался преимущественно как отдельная область знаний, требующая высокой теоретической подготовки (математическое моделирование, алгоритмика, машинное обучение), то сегодня он стремительно превращается в сквозную компетенцию, или soft skill, необходимую практически каждому специалисту XXI в.

Не менее важным направлением институционализации является разработка нормативных документов, регулирующих использование ИИ в образовательных учреждениях. Например, методические рекомендации по применению генеративных сервисов в школах и вузах должны учитывать не только технические возможности, но и педагогическую целесообразность, этические стандарты, требования к защите данных. В этой связи целесообразно, чтобы рабочие группы по подготовке таких документов включали экспертов, чьи исследования и практические разработки находят отражение на ведущих научных форумах. Это позволит преодолеть разрыв между академической наукой и прикладными потребностями, а также обеспечит нормативное сопровождение, соответствующее реальным вызовам.

Таким образом, задача белорусской образовательной и научной системы состоит в переходе от стихийного распространения инноваций к их институциональной интеграции. Для этого необходимы:

– создание централизованных структур регламентирования ИИ;

– разработка нормативных документов с участием академических и практических экспертов;

– обеспечение стратегического финансирования и координации усилий.

Для сопоставления показателен опыт Российской Федерации. В рамках реализации Национальной стратегии развития искусственного интеллекта на период до 2030 года создан Национальный фонд искусственного интеллекта (НФИИ), являющийся ассоциацией организаций и специалистов, которая координирует деятельность в сфере ИИ. В его структуру в рамках общей стратегии регулирования ИИ в России входят следующие экспертные и специализированные комитеты ([6], а также QR-код):



– Комитет по искусственному интеллекту в образовании и науке;

– Комитет по искусственному интеллекту в здравоохранении;

– Комитет по искусственному интеллекту в строительстве;

– Комитет по когнитивной безопасности;

– Комитет по международному сотрудничеству в области искусственного интеллекта.

На государственном и законодательном уровнях России действуют или формируются следующие структуры:

– Российский комитет по вопросам этики искусственного интеллекта (при Комиссии РФ по делам ЮНЕСКО), который отвечает за выработку подходов к международному сотрудничеству в области этики ИИ;

– рабочая группа по совершенствованию нормативного правового регулирования (при Комитете Совета Федерации);

– рабочая группа по разработке законов о применении искусственного интеллекта (при Госдуме), которая занимается законодательным обеспечением и отраслевым применением ИИ;

– технический комитет по стандартизации 164 «Искусственный интеллект», отвечающий за разработку национальных стандартов, терминологии и правил применения ИИ и др.

Такая система отражает трансдисциплинарный подход, где регулирование и развитие ИИ охватывают как технические, так и этические, правовые, а также узкоотраслевые аспекты. Эти структуры обеспечивают стратегическую координацию и экспертное сопровождение проектов, что позволяет минимизировать дублирование усилий и концентрировать ресурсы на наиболее перспективных направлениях. В результате инновации получают институциональную опору: отдельные инициативы трансформируются в национальные программы, которые охватывают и научные исследования, и подготовку кадров, и внедрение технологий в ключевые отрасли. Такой системный подход демонстрирует, что без развитых институциональных механизмов инновации в сфере ИИ не смогут обрести устойчивость и масштаб. Именно такие шаги позволят превратить отдельные инициативы в единую экосистему, способную готовить специалистов нового

поколения и обеспечивать долгосрочное развитие национального потенциала в сфере ИИ.

Сравнивая белорусский и российский опыт, можно выделить несколько принципиальных различий. Беларусь в аспекте внедрения ИИ развивается по пути «снизу вверх» – от отдельных практик к возможной консолидации; Россия – «сверху вниз», через стратегические документы и централизованные структуры. В России задействованы значительные финансовые и организационные механизмы, в Беларуси пока преобладает энтузиазм сообществ и отдельных образовательных учреждений. При этом белорусская модель обеспечивает гибкость и адаптивность, но страдает от фрагментации; российская модель гарантирует системность, но может быть менее чувствительной к локальным потребностям.

Для Беларуси представляется целесообразным адаптировать элементы российской модели с учетом национального контекста:

– создание централизованного фонда как центра компетенций, обеспечивающего координацию;

– дальнейшая разработка нормативных документов и методических рекомендаций по применению ИИ в образовании;

– формирование системы экспертизы и оценки качества внедряемых решений.

Сравнение показывает, что устойчивое развитие ИИ в Беларуси невозможно без институционализации и консолидации усилий. Локальные инициативы создают основу, но именно национальная политика может превратить их в единую экосистему, способную конкурировать на международном уровне. Таким образом, вопрос не только в том, какие технологии внедрять, но и кто будет готовить специалистов для работы с ними. Ответом на этот вызов может стать создание консорциумов – гибких и многоуровневых объединений, способных связать академическую науку, образование, бизнес и государство в единую систему подготовки кадров для цифровой эпохи.

Заключение

В ходе анализа перформативных аспектов внедрения ИИ в образовательное пространство Беларуси выявлен существенный разрыв между академической повесткой, ориентированной преимущественно на теоретические исследования, и стихийными практиками педагогов и студентов, которые уже активно применяют ИИ-сервисы в повседневном обучении.

Классифицированы ключевые барьеры процесса внедрения: технологические (быстрое устаревание учебных программ и инструментов), концептуально-дидактические (нехватка преподавателей с актуальными компетенциями, дисбаланс между теорией и практикой) и отраслевые (рассин-

хронизация между академическим образованием и потребностями рынка труда). Их фиксация обеспечивает понимание того, что без устранения этих структурных дисбалансов внедрение ИИ будет носить фрагментарный характер и не приведет к формированию устойчивой национальной образовательной модели в рамках концепций Университет 4.0 и Университет 5.0.

Определены направления институционализации инноваций, включающие создание трансдисциплинарных консорциумов, разработку образовательных стандартов нового поколения и формирование национальной программы подготовки кадров. Эти меры позволяют перевести процесс внедрения ИИ из стадии стихийных практик в системное развитие, обеспечивая согласованность академической подготовки с запросами общества и экономики, а также формируя базу для долгосрочной устойчивости образовательной экосистемы.

Сопоставлены белорусская и российская модели развития в сфере ИИ, демонстрирующие противоположные траектории: Беларусь развивается «снизу вверх» за счет локальных инициатив, тогда как Россия движется «сверху вниз» через централизованные механизмы и стратегические документы. Такое сравнение позволяет не только зафиксировать различия в подходах, но и выявить возможности их синтеза: адаптация отдельных элементов российской модели в белорусский контекст может обеспечить сочетание гибкости и адаптивности локальных практик с системностью и устойчивостью институциональных решений, что критически важно для формирования целостной образовательной экосистемы ИИ.

Список использованных источников

1. Цыркун, И. И. Система инновационной подготовки специалистов гуманитарной сферы / И. И. Цыркун. – Минск : Тэхналогія, 2000. – 326 с.

2. Первая выставка-форум IT-академграда «Искусственный интеллект в Беларуси» : сб. докл., Минск, 13–14 окт. 2022 г. – Минск : ОИПИ НАН Беларуси, 2022. – 122 с.

3. II Форум IT-Академграда «Искусственный интеллект в Беларуси» : доклады, Минск, 12–13 окт. 2023 г. – Минск : ОИПИ НАН Беларуси, 2023. – 276 с.

4. III Форум IT-Академграда «Искусственный интеллект в Беларуси» : доклады, Минск, 10–11 окт. 2024 г. – Минск : ОИПИ НАН Беларуси, 2024. – 352 с.

5. Пунчик, В. Н. Терминологический редизайн педагогики в условиях цифровой трансформации / В. Н. Пунчик // Научные труды Республиканского института высшей школы. – 2024. – Вып. 24. – С. 167–175.

6. Национальный фонд искусственного интеллекта. – URL: https://nfai.ru (дата обращения: 03.10.2025).

УДК 004.852

Анализ данных для принятия решений в конструкторской подготовке

И. А. Резниченко✉, **И. В. Акиншева**
Могилевский государственный университет
имени А. А. Кулешова, Беларусь
E-mail: reznichenkoivan907@gmail.com

Введение

Современное производство энергоэффективных электродвигателей сталкивается с необходимостью строгого соответствия отечественным и международным стандартам качества, что обусловлено их значимостью как ключевых компонентов промышленных систем и экспортируемой продукции. В свою очередь, повышение конкурентоспособности таких изделий требует внедрения автоматизированных систем, способных не только собирать и обрабатывать данные испытаний, но и предоставлять аналитические выводы для совершенствования конструкторских решений [1].

Актуальность представленной работы заключается в создании системы анализа данных, которая интегрирует результаты испытаний, формирует структурированное хранилище и закладывает основу для применения методов машинного обучения, способствующих оптимизации проектирования и конструирования электродвигателей. Целью исследования является разработка концепции аналитической системы, способной трансформировать данные испытаний в основу для принятия решений в конструкторской подготовке. Для ее достижения необходимо решить следующие задачи:

1) описать протокол испытаний двигателя, включая его структуру и ключевые измеряемые параметры;

2) описать методику расчета технических характеристик двигателя, используемых при формировании протокола;

3) разработать проект хранилища данных, обеспечивающего сбор, хранение и анализ информации с целью формирования аналитики.

Основная часть

Протокол испытаний выступает центральным документом, фиксирующим результаты тестирования электродвигателя, и служит основанием для принятия решений о его пригодности. Его структура строго регламентирована и включает идентификационные данные образца, перечень нор-

мативных документов (ГОСТ, СТБ, IEC), которые устанавливают методы испытаний, таблицу измеренных параметров и заключение о соответствии [2]. В таблицу результатов занесены как нормированные, так и фактические значения, что позволяет проводить прямое сопоставление. Двигатель считается пригодным, если все его измеренные параметры попадают в нормируемые диапазоны, установленные стандартами [3]. Например, для определенного типа двигателей установлены следующие диапазоны: ток холостого хода I_0 – от 2,11 до 236 А, минимальный ток короткого замыкания I_{kmin} – от 2,41 до 4,01 А, максимальный ток I_{kmax} – от 2,37 до 3,95 А, мощность холостого хода P_0 – от 2,06 до 242 Вт, минимальная мощность короткого замыкания P_{kmin} – от 321 до 534 Вт и максимальная P_{kmax} – от 306 до 510 Вт. Выход за пределы этих диапазонов ведет к отбраковке образца, однако повторяющиеся отклонения в одной и той же зоне могут, например, указывать на системную ошибку в конструкции, которая выявляется только при системном анализе.

Расчет характеристик двигателя основан на данных, поступающих в режиме реального времени с первичных преобразователей, таких как датчики тока, напряжения и момента, в частности системы M40-X с декодером T45 и анализаторы Elspec G4420. Эти сигналы подвергаются предварительной обработке, прежде чем быть использованными для формирования итоговых значений, включаемых в протокол. Интерфейсное окно оператора испытания механической характеристики приведено на рис. 1.

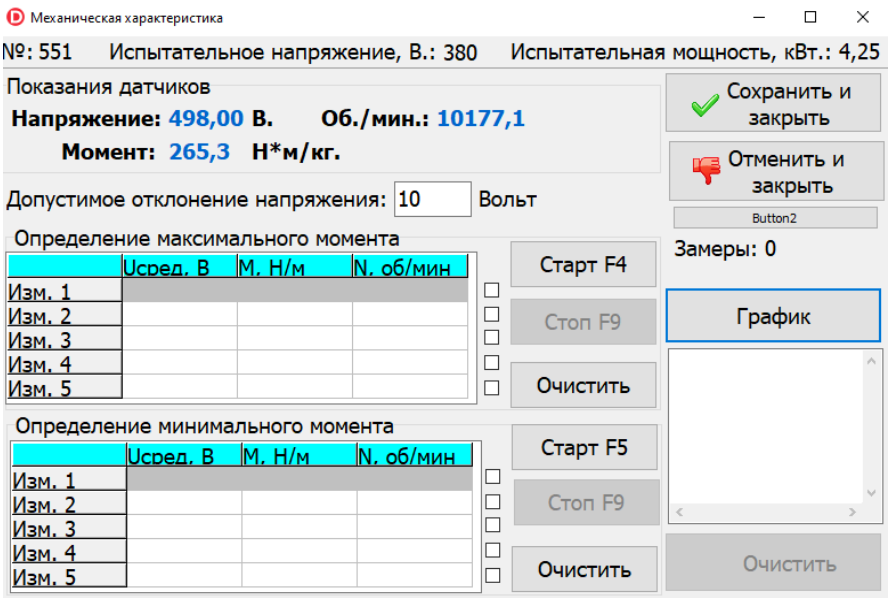


Рис. 1. Окно испытания механической характеристики

Обработка сигналов, получаемых от первичных преобразователей, состоит из применения алгоритмов фильтрации, направленных на устране-

ние шумов, которые могут быть как постоянными, так и случайными. В частности, используются медианный фильтр и модифицированный фильтр Калмана, причем последний реализуется итерационно с учетом параметров, таких как интервал дискретизации, среднеквадратичное отклонение процесса и дисперсия шума.

Ошибка прогнозирования r, появляющаяся из-за шума, находится следующим образом:

$$r = \frac{4+\frac{\sigma_\omega \cdot T^2}{\sigma_\upsilon} - \sqrt{8\cdot\frac{\sigma_\omega \cdot T^2}{\sigma_\upsilon} + \left(\frac{\sigma_\omega \cdot T^2}{\sigma_\upsilon}\right)^2}}{4},$$

где T – интервал дискретизации; σ_ω^2 – среднеквадратичное отклонение процесса; σ_υ^2 – дисперсия шума.

Эти алгоритмы, в свою очередь, позволяют минимизировать среднеквадратичную ошибку, обеспечивая высокую точность данных, которые затем применяются для построения графика замера минимального момента двигателя (рис. 2).



Рис. 2. График замера минимального момента

Точность измерений имеет решающее значение для корректного анализа характеристик, а некий уровень погрешности, возникающий из-за шумов, устраняется благодаря тщательной настройке параметров фильтрации.

Для расчета технических характеристик, таких как ток, напряжение и механические параметры, применяются алгоритмы усреднения, основанные на методе скользящего среднего и экспоненциального бегущего среднего. Эти методы, реализованные в программном обеспечении, позволяют оператору получать усредненные значения сигналов в реальном времени, что отражается в интерфейсном окне, где отображается линия тренда ме-

ханической характеристики с возможностью выбора опорных точек для вычисления момента в процессе разгона двигателя. В свою очередь, поддержание напряжения с точностью до 1 В, обеспечиваемое программно-аппаратным комплексом с использованием микроконтроллера ATMega328 и реле, исключает деформацию характеристик, что крайне важно для достоверности результатов в протоколе.

Полученные данные, обработанные с применением описанных алгоритмов, автоматически заносятся в электронный паспорт электродвигателя, который формируется в соответствии с международными стандартами, а также используются для построения графиков, например при определении минимального момента в процессе разгона, что невозможно сделать по сырым данным. Таким образом, итоговые значения в протоколе – это не просто замеры, а результат алгоритмической обработки, обеспечивающей высокую точность и воспроизводимость.

Для эффективного анализа данных, накопленных в ходе испытаний, требуется надежное и масштабируемое хранилище, способное обеспечить целостность информации и поддержать сложные запросы. При выборе системы управления базами данных (СУБД) рассматриваются два основных варианта – Oracle и PostgreSQL. Обе системы обладают высокой производительностью, поддержкой сетевого доступа и развитыми средствами резервного копирования. Однако их отличия становятся существенными в условиях промышленной эксплуатации.

PostgreSQL – это мощная открытая СУБД с широкими возможностями расширения и гибкой моделью репликации. Тем не менее ее производительность в высоконагруженных сценариях может уступать коммерческим решениям при интенсивной записи данных с нескольких стендов в реальном времени. Oracle, в свою очередь, демонстрирует превосходную стабильность, высокую скорость обработки транзакций и встроенные механизмы кластеризации, что критично для обеспечения непрерывности работы. Кроме того, Oracle предоставляет развитые инструменты для мониторинга, резервного копирования и восстановления, что минимизирует риски потери данных при сбоях оборудования [4].

Исходя из этих факторов, выбор делается в пользу Oracle как более надежного и производительного решения для промышленной среды. Хранилище данных проектируется как централизованная сетевая база, расположенная на защищенном сервере предприятия. Оно включает таблицы для хранения идентификационных данных двигателей (тип, серийный номер, мощность, напряжение), параметров испытаний (режим, длительность, условия), а также массивы необработанных сигналов и итоговых рассчитанных характеристик. Все данные сохраняются в полном объеме, что позволяет в любой момент провести повторный анализ с использова-

нием новых алгоритмов. Связи между таблицами обеспечивают целостность данных и удобство построения аналитических отчетов.

В данный момент анализ в протоколе испытаний сводится к сравнению измеренных значений с нормативными диапазонами. Основное внимание уделяется таким параметрам, как ток холостого хода I_0, токи короткого замыкания I_{kmin} и I_{kmax}, а также соответствующие мощности P_0, P_{kmin} и P_{kmax}. Сравнение происходит по принципу «входит/не входит» в допустимый диапазон, что эффективно для отбраковки, но может не разрешить выявить системные тенденции. Аналитическая система, построенная на основе хранилища данных, позволит выйти за рамки единичных испытаний и провести сравнение с историческими данными по аналогичным моделям, выявляя повторяющиеся отклонения, которые в совокупности могут указывать на необходимость корректировки конструкторской документации.

Переход от простого сравнения к предиктивному анализу требует применения методов машинного обучения. Учитывая структуру данных: числовые временные ряды и табличные характеристики, перспективными являются алгоритмы, способные работать с многомерными признаками и выявлять сложные зависимости. Линейная регрессия может служить базовой моделью для прогнозирования значений параметров по входным данным двигателя на основе исторических данных, однако ее возможности ограничены линейностью. Более гибким решением являются ансамблевые методы, такие как случайный лес (random forest) или градиентный бустинг (gradient boosting), которые эффективно справляются с нелинейностями и способны ранжировать признаки по их важности для предсказания, например величины I_0 [5].

Внедрение машинного обучения включает несколько этапов: обучение моделей на исторических данных, валидацию на тестовых выборках, оценку точности и последующую интеграцию в производственный цикл. Например, после обучения модели на данных прошлых испытаний ее можно использовать для автоматического анализа новых протоколов, что сократит время на обработку и снизит риск человеческих ошибок.

Заключение

Разработанная архитектура аналитической системы открывает возможности для перехода от реактивного контроля качества к проактивному управлению конструкторской подготовкой. Централизованное хранилище данных на базе Oracle обеспечивает надежность и масштабируемость, позволяя накапливать полную историю испытаний. Алгоритмы обработки сигналов гарантируют высокую точность измерений, что является фундаментом для достоверного анализа. Внедрение машинного обучения, осно-

ванное на накопленных данных, позволит выявить скрытые закономерности и системные отклонения, которые невозможно обнаружить при традиционном подходе. Скорость обучения моделей будет возрастать по мере роста объема данных, что сделает систему более точной. В конечном счете это обеспечит не только улучшение конструкции двигателей, но и повышение конкурентоспособности продукции на мировом рынке.

Список использованных источников

1. Акиншева И. В. Автоматизация расчета испытательных характеристик электродвигателей / И. В. Акиншева, С. В. Микулич // Приборостроение-2022: материалы 15-й Междунар. науч.-техн. конф., Минск, 16–18 нояб. 2022 г. / Белорусский национальный технический университет. – С. 72–73.

2. Somvanshi, U. Quality Control in Automobile Manufacturing Industries / U. Somvanshi // International Journal for Research in Applied Science and Engineering Technology. – 2020. – Vol. 8. – Pp. 740–750.

3. Pacana, A. Analysis of quality control efficiency in the automotive industry / A. Pacana, K. Czerwińska, R. Dwornicka // Transportation Research Procedia. – 2021. – Vol. 55. – P. 691–698.

4. Reddy Kavuluri, H. V. PostgreSQL vs. Oracle: A comparative study of performance, scalability, and enterprise adoption / H. V. Reddy Kavuluri // International Journal of Emerging Trends in Computer Science and Information Technology. – 2024. – Vol. 5, no. 2. – P. 33–41.

5. Введение в статистическое обучение с примерами на языке R / Г. Джеймс, Д. Уиттон, Т. Хасти, Р. Тибширани; пер. с англ. С. Э. Мастицкого. – М. : ДМК Пресс, 2016. – 450 с.

УДК 519.714

Экспериментальное сравнение алгоритмов схемной реализации функций многозначной логики

В. И. Романов, П. Н. Бибило✉
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: bibilo@newman.bas-net.by

Введение

Автоматизированное проектирование аппаратуры цифровых систем привело к созданию разнообразных средств технологически независимой оптимизации, предварительно выполняемой непосредственно перед логическим синтезом схем. Система *FLC-2* [1], основанная на применении продукционно-фреймовой модели представления знаний [2] и предназначенная для технологически-независимой логической оптимизации функционально-структурных описаний цифровых блоков СБИС, была пополнена новыми программами оптимизации представлений функций многозначной логики. Это позволило проводить логическую оптимизацию различными процедурами, использующими как BDD-представления (*BDD – Binary Decision Diagram*, бинарная диаграмма решений) систем булевых функций, так и MDD-представления функций многозначной логики (*MDD – Multi-valued Decision Diagram,* многозначная диаграмма решений) [3]. В данной работе сообщается о результатах экспериментального сравнения алгоритмов схемной реализации в заказных СБИС функций многозначной логики.

Исходные данные для экспериментов

Таблицы истинности, задающие булевы функции арифметических (и модулярных) сумматоров и умножителей, были получены выполнением следующих этапов:

– составлены VHDL-модели описания функционирования устройств;

– для модели каждого устройства проведено моделирование на всех наборах соответствующего булева пространства;

– по результатам моделирования каждого устройства составлена таблица истинности соответствующей системы булевых функций.

Для арифметических 2-операндных сумматоров, умножителей и линейных функций это были системы полностью определенных булевых функций. Для модулярных 2- и 3-операндных устройств соответствующие

системы булевых функций были частично определенными (частичными) – для некоторых наборов значений аргументов значения систем функций задавалось вектором, состоящим только из неопределенных значений «–».

Эксперименты проводились на трех потоках примеров:

Поток 1. Схемы для реализации линейной функции и функций арифметического сложения и умножения.

Поток 2. Схемы для 2-операндных модулярных устройств (сумматоров, умножителей).

Поток 3. Схемы для 3-операндных модулярных устройств (сумматоров, умножителей).

Приведем обозначения устройств для арифметического сложения и умножения (поток 1 примеров). Например, Adder_6_2 – это 2-операндный *сумматор* для сложения ($y=a+b$) целых неотрицательных чисел a, b, заданных 6-разрядным двоичным кодом, Multiplier_8_2 – это 2-операндный *умножитель* ($y=ab$) целых неотрицательных чисел, заданных 8-разрядным двоичным кодом. Схема с именем Linf_3_3 задает 3-операндную *линейную* функцию ($y=ab+c$), каждый из операндов a, b, c задает неотрицательное целое число, заданное 3-разрядным двоичным кодом. Подобным образом описываются сумматоры и умножители (далее такие описания используются в эксперименте 4).

Имена и значения параметров сгенерированных функциональных описаний *модулярных* устройств (потоки 2 и 3 примеров): через Sum_p_k, Mult_p_k обозначено k-операндное (k=2, 3) устройство по модулю p, сумматор (Sum) и умножитель (Mult) соответственно. Например, через Sum_15_3 обозначен 3-операндный сумматор по модулю 15.

Оптимизационные алгоритмы

Технологически независимая оптимизация в экспериментах выполнялась различными оптимизационными алгоритмами, основу которых составляют две программы – *BDD_Builder* и *MDD_Builder* [2].

Приведем список условных обозначений оптимизационных алгоритмов, использованных в экспериментах схемной функции многозначной логики:

BDD – минимизация BDD-представлений систем булевых функций.

MDD-1 – MDD-минимизация многозначной функции, входные переменные a, b, c и арифметическая функция y (поток 1) являются многозначными.

MDD-2 – MDD-минимизация полностью определенной многозначной функции $y=y(a, b)$, пары (a_i, b_i) булевых переменных, образованные одноименными разрядами (a_i, b_i), $i=n, n\text{-}1,\ldots,1$, складываемых либо умножаемых чисел интерпретируются как 4-значные аргументы x_i. Построение MDD ведется по перестановке $<x_n, \ldots, x_1>$, т. е. начиная с младших разря-

дов складываемых либо умножаемых чисел. Выходная функция является многозначной. Процедура *MDD-2* применяется для потоков 1 и 2 примеров схем. Для линейных функций $y=ab+c$ многозначные переменные образуются из троек (a_i, b_i, c_i), $i=n, n-1,\ldots,1$, булевых переменных.

MDD-3 – MDD-минимизация, аналогичная *MDD-2*, однако вместо одной многозначной выходной функции в *MDD-3* рассматриваются булевы выходы, что соответствует случаю системы конечных предикатов.

MDD-4 – MDD-минимизация системы конечных предикатов, полученной из системы булевых функций перебором 1000 случайных перестановок множества аргументов, для каждой перестановки *пары* соседних аргументов рассматриваются как 4-значные аргументы x_i, после чего строится MDD для фиксированной перестановки многозначных аргументов x_i и осуществляется переход к логическим уравнениям. Для логических уравнений подсчитывается суммарное число *литералов* (булевых переменных либо их инверсий) и выбирается та перестановка, для которой число литералов является наименьшим.

MDD-5 – MDD-минимизация системы конечных предикатов, полученной из системы булевых функций перебором 1000 случайных перестановок множества аргументов, для каждой перестановки *тройки* соседних аргументов рассматриваются как 8-значные аргументы x_i, после чего строится MDD для многозначных аргументов x_i и осуществляется переход к логическим уравнениям. Для логических уравнений подсчитывается суммарное число литералов и выбирается та перестановка, для которой число литералов является наименьшим.

Эксперименты

Каждый из перечисленных далее 10 экспериментов состоял из трех этапов:

1. Выполнение оптимизационной процедуры – булевой (*BDD*) либо многозначной (*MDD-1 – MDD-5*) минимизации и получение логических уравнений, задающих BDD либо MDD. Эксперименты различаются только этапом 1, на котором для различных потоков исходных данных выполняются различные оптимизационные процедуры.

2. Конвертация логических уравнений в VHDL-описание.

3. Синтез логической схемы. VHDL-описание подается на вход синтезатора LeonardoSpectrum – выполняется синтез логической схемы в библиотеке проектирования заказных СБИС. Библиотекой синтеза во всех экспериментах являлась одна и та же библиотека проектирования заказных цифровых КМОП СБИС [4].

Результаты экспериментов

Результаты экспериментов 1 – 4 для первого потока примеров приведены в табл. 1, экспериментов 5 – 8 для 2-операндных модулярных устройств (поток 2 примеров) – в табл. 2, экспериментов 9 и 10 для 3-операндных модулярных устройств (поток 3 примеров) – в табл. 3. Во всех таблицах с результатами экспериментов L – число литералов в логических уравнениях; S – суммарная площадь элементов схемы в условных единицах; t – задержка схемы (нс). Символом * выделены (в рамках одной таблицы) лучшие решения для испытываемого примера – меньшее значение параметра площади схемы, меньшее значение задержки, меньшее число литералов в соответствующем функциональном описании (логических уравнениях).

Таблица 1

Схемная реализация арифметических функций

Имя схемы (поток 1)	*Булева минимизация*			*Многозначная минимизация*					
	Эксперимент 1 *BDD*			Эксперимент 2 *MDD-1*			Эксперимент 3 *MDD-2*		
	L	*S*	*t*	*L*	*S*	*t*	*L*	*S*	*t*
Adder_2_2	*30	*3543	2.54	77	*3354	*2.15	60	3739	2.80
Adder_3_2	*60	*6428	3.61	496	17 404	3.12	246	6590	*2.10
Adder_4_2	*90	*7661	*3.90	2923	63 517	6.35	760	9581	3.69
Adder_5_2	*120	*11 807	*3.94	16 426	187 912	9.32	2034	16 885	3.97
Adder_6_2	*150	*14 101	*4.62	88 953	454 681	15.98	5012	22 900	5.25
Adder_7_2	*180	*17 030	*5.52	466 488	-	-	11 742	27 521	6.30
Adder_8_2	*210	*19 826	7.76	2 378 999	-	-	26 640	34 256	*7.47
Multiplier_2_2	*30	3214	2.64	5	*2332	*1.76	52	*2332	*1.76
Multiplier_3_2	*144	*14 754	*3.27	457	22 247	3.44	447	18 816	3.46
Multiplier_4_2	*592	72 272	*5.17	3182	75 380	7.29	2676	*70 074	5.29
Multiplier_5_2	*2135	258 360	*7.61	20 183	264 074	9.93	14 303	*243 249	9.01
Multiplier_6_2	*6318	945 408	11.98	118 170	848 154	15.09	71 287	*800 317	*11.73
Multiplier_7_2	*22 827	3 246 745	*13.79	658 586	-	-	341 919	*2 722 376	13.98
Multiplier_8_2	*66 138	8 358 131	*17.32	3 509 482	-	-	1 592 983	-	-
Linf_2_3	*119	*12 326	4.47	221	13 827	*2.97	248	15 256	3.22
Linf_3_3	*643	71 296	*5.20	2900	*67 395	5.35	4933	80 994	5.96
Linf_4_3	*3074	342 406	8.68	28 483	*284 083	9.60	54 219	307 949	*8.42
Linf_5_3	*12 981	1 578 431	*12.09	291 460	-	-	553 725	*1 181 565	12.88
Лучших решений	18	9	11	0	4	3	0	6	5

Таблица 2

Схемная реализация модулярных 2-операндных умножителей и сумматоров

Имя схемы (поток 2)	*Булева минимизация*			*Многозначная минимизация*		
	Эксперимент 5 *BDD*			Эксперимент 6 *MDD-2*		
	L	*S*	*t*	*L*	*S*	*t*
Mult_5_2	*84	*10 474	3.04	24	10 596	2.90
Mult_7_2	*138	*13 554	2.75	40	15 440	3.50
Mult_9_2	*281	*31 259	5.08	86	37 202	5.28
Mult_13_2	*512	66 285	5.69	264	65 152	*5.24
Mult_15_2	*364	*43 267	*4.84	536	77 752	6.01
Mult_17_2	*1 027	114 848	*7.22	824	106 969	*7.22
Mult_29_2	*2 181	280 557	9.02	8 076	308 022	*7.45
Mult_33_2	*2 603	326 106	*8.56	3 138	351 350	9.45
Mult_41_2	*5 120	675 225	*10.05	21 448	608 761	10.14
Mult_63_2	*3 228	*395 594	11.61	44 352	1 215 720	11.01
Sum_5_2	95	11 779	*3.36	34	12 164	4.99
Sum_7_2	*144	16 121	3.02	235	*13 961	*2.16
Sum_9_2	*266	30 160	*4.30	554	31 639	4.40
Sum_15_2	*390	42 614	*4.46	979	46 448	5.12
Sum_17_2	*507	*47 095	*4.97	1 802	54 851	6.74
Sum_19_2	*554	59 589	*5.12	2 059	65 638	5.99
Sum_23_2	*583	*61 826	*5.06	2 425	64 159	6.68
Sum_25_2	*643	65 933	*5.65	2 797	64 851	6.18
Sum_27_2	*722	80 012	*6.76	2 971	76 714	7.72
Sum_29_2	*729	74 599	*5.36	3 097	*68 656	6.91
Sum_31_2	*667	71 162	6.91	3 019	70 565	*6.49
Sum_37_2	*939	*82 405	*6.81	5 883	95 859	8.35
Sum_59_2	*1073	111 717	*8.10	8 419	*107 521	9.87
Sum_61_2	*1148	111 193	8.49	8 431	102 315	*7.39
Лучших решений	23	9	5	0	3	7

Окончание табл. 2

Имя схемы (поток 2)	*Многозначная минимизация*					
	Эксперимент 7 *MDD-3*			Эксперимент 8 *MDD-4*		
	L	*S*	*t*	*L*	*S*	*t*
Mult_5_2	95	10 596	*2.90	86	10 596	*2.90
Mult_7_2	170	15 440	3.50	152	17 164	*2.71
Mult_9_2	407	35 333	4.38	317	34 122	*3.97
Mult_13_2	730	66 346	5.33	652	*61 123	6.10
Mult_15_2	858	80 106	6.98	412	50 839	5.12
Mult_17_2	1 268	111 260	7.45	1 118	*97 377	7.32
Mult_29_2	4 243	311 598	8.56	3 112	*259 141	9.07
Mult_33_2	4 631	347 249	9.62	3 379	*314 935	9.83
Mult_41_2	8 173	613 426	10.77	6 763	*583 300	10.18
Mult_63_2	16 650	1 205 548	10.67	4 386	405 337	*9.14
Sum_5_2	107	12 164	4.99	*93	13 593	4.65
Sum_7_2	183	*13 961	*2.16	183	*13 961	*2.16
Sum_9_2	320	31 639	4.40	308	*28 709	4.94
Sum_15_2	541	*42 084	5.75	544	43 920	5.01
Sum_17_2	632	53 875	6.82	650	57 971	6.21
Sum_19_2	799	61 977	6.95	759	*59 120	6.90
Sum_23_2	853	64 443	6.25	865	68 746	6.60
Sum_25_2	848	64 393	5.75	860	*62 714	5.67
Sum_27_2	1 030	*75 553	7.83	1 042	76 904	7.66
Sum_29_2	988	74 560	6.50	1 000	74 510	7.09
Sum_31_2	1 107	*67 680	6.61	1 119	77 496	8.14
Sum_37_2	1 312	96 037	7.95	1 602	111 025	7.46
Sum_59_2	1 770	107 945	9.86	2 303	138 016	10.78
Sum_61_2	1 686	*100 998	8.94	2 348	138 183	9.14
Лучших решений	0	5	2	1	9	5

Таблица 3

Схемная реализация модулярных 3-операндных умножителей и сумматоров

Имя схемы (поток 3)	*Булева минимизация*		*Многозначная минимизация*	
	Эксперимент 9 *BDD*		Эксперимент 10 *MDD-5*	
	S	*t*	*S*	*t*
Mult_3_3	*5 586	2.66	5 669	*2.43
Mult_5_3	30 076	4.46	*20 992	*4.31
Mult_7_3	*28 837	*4.35	29 836	4.67
Mult_9_3	109 876	7.72	*109 212	*6.87
Mult_13_3	552 342	10.51	313 178	*10.42
Mult_15_3	*119 139	*8.37	155 319	8.50
Mult_17_3	1 047 907	13.16	*645 321	*11.34
Mult_29_3	2 967 388	15.70	*2 215 171	*15.63
Sum_3_3	*9 559	3.46	11 009	*3.22
Sum_5_3	*24 485	5.73	27 682	*5.19
Sum_7_3	*42 877	*6.58	61 134	7.43
Sum_9_3	*61 296	*6.95	114 881	7.01
Sum_13_3	*107 794	9.47	200 908	*8.48
Sum_15_3	*127 285	*9.64	276 400	12.36
Sum_17_3	*151 374	*9.81	225 080	10.72
Sum_29_3	1 027 250	14.10	*829 489	*13.77
Лучших решений	10	6	6	10

Поток 1. Анализируя табл. 1 с результатами экспериментов для потока 1 арифметических функций, легко установить, что для схем сложения эффективнее применение булевой BDD-минимизации. Для умножителей и линейных функций более эффективно применение многозначной MDD-минимизации, объявляя многозначными переменными пары либо тройки одноименных булевых переменных (разрядов представления аргументов) в системах булевых функций. Проведение дальнейших экспериментов показало, что замена синтезируемых VHDL-операций арифметических функций таблицами истинности является нецелесообразной, так как результаты синтеза по синтезируемым алгоритмическим описаниям дают схемы меньшей площади.

Поток 2. Для 2-операндных модулярных схем (табл. 2) перебор случайных перестановок булевых переменных в системах булевых функций

дает наибольший эффект по числу лучших результатов (по площади) для 2-операндных модулярных умножителей, заданных системами конечных предикатов – это результаты эксперимента 7. Другие процедуры многозначной MDD-минимизации позволяют получать лучшие результаты для 2-операндных модулярных сумматоров.

Поток 3. При схемной реализации 3-операндных модулярных схем (табл. 3) чаще всего лучшие результаты по площади (10 примеров) можно получить, применяя BDD-минимизацию. Однако MDD-минимизация дает чаще (также 10 примеров) лучшие задержки схем по сравнению с BDD-минимизацией, которая приводит к лучшим задержкам схем только в шести случаях.

Заключение

Результаты экспериментов показывают эффективность применения при синтезе заказных СБИС различных алгоритмов булевой и многозначной технологически независимой оптимизации. Можно отметить то, что для арифметических сумматоров BDD-оптимизация позволяет получать схемы, близкие по площади к схемам, получаемым по синтезируемым арифметическим операциям сложения в промышленных синтезаторах схем, что показывает конкурентоспособность системы *FLC-2* в области логического проектирования блоков комбинационной логики СБИС. Другие эксперименты, проведенные для схем управляющей логики, показали, что имеется много возможностей для разработки различных алгоритмов перехода от моделей систем булевых функций к моделям многозначных функций (и систем) для выполнения многозначной технологически независимой оптимизации.

Список использованных источников

1. Бибило, П. Н. Система логической оптимизации функционально-структурных описаний цифровых устройств на основе продукционно-фреймовой модели представления знаний / П. Н. Бибило, В. И. Романов // Проблемы разработки перспективных микро- и наноэлектронных систем : сб. тр. / под общ. ред. акад. РАН А. Л. Стемпковского. – М. : ИППМ РАН, 2020. – № 4. – С. 9–16.

2. Бибило, П. Н. Логическое проектирование дискретных устройств с использованием продукционно-фреймовой модели представления знаний / П. Н. Бибило, В. И. Романов. – Минск : Беларус. навука, 2011. – 279 с.

3. Бибило, П. Н. Технологически независимая оптимизация функций многозначной логики / П. Н. Бибило, В. И. Романов (в наст. сборнике).

4. Бибило, П. Н. Бинарные диаграммы решений в логическом проектировании / П. Н. Бибило. – М. : Ленанд, 2024. – 560 с.

УДК [004.934+681.84.081]-028.17:001.32(476)

Технологии синтеза речи в аудиогидах: проблемы и перспективы

М. М. Слесарева, Е. С. Зеновко, А. А. Макогон, Г. С. Кухаревич, В. А. Хохлов
Объединенный институт проблем информатики Национальной академии наук Беларуси, Минск
E-mail: secretary.tcii@newman.bas-net.by

Введение

Создание аудиогидов является важным шагом в повышении доступности и улучшении опыта для посетителей различных учреждений, таких как музеи, исторические места, парки, природные заповедники, культурные центры и другие туристические объекты. Аудиогиды предоставляют посетителям возможность получить более подробную информацию о представленных экспонатах, исторических событиях или природных особенностях территории, что способствует лучшему восприятию окружающей среды.

Одним из главных преимуществ аудиогидов является их доступность для людей с ограниченными возможностями: слабовидящих или людей с нарушениями слуха. Аудиогиды позволяют получить равный доступ к информации, которая может быть недоступна или труднодоступна в других форматах. Кроме того, аудиогиды могут быть полезны для туристов, которые предпочитают самостоятельно исследовать новые места. Это особенно актуально для людей, которые не имеют возможности или не хотят присоединяться к групповым экскурсиям. Аудиогиды также могут быть использованы в образовательных целях, позволяя студентам и школьникам получить более глубокое понимание исторических событий, культурных традиций или научных концепций. Это может быть особенно полезно для детей, которые заинтересованы в обучении через аудиоформат.

Современные аудиогиды претерпели значительные изменения с развитием технологий, превратившись из простых аудиорекордеров с заранее записанной информацией в сложные системы, интегрированные с различными технологиями [1]. Интеграция с мобильными приложениями, GPS-навигацией, технологиями синтеза и распознавания речи стала ключевым фактором, который расширил их функциональность и сделал информацию более интерактивной и персонализированной. Интеграция аудиогидов с мобильными приложениями в первую очередь позволяет улучшить пользова-

тельский опыт: пользователи могут получить доступ к аудиогидам непосредственно на своих смартфонах или планшетах, что делает их использование более удобным. Мобильные приложения могут дополнить аудиогиды такими функциями, как тестирование знаний, интерактивные задания, и возможностью делиться впечатлениями в социальных сетях. Важным аспектом использования аудиогидов на смартфонах является сбор данных разработчиками о поведении пользователей, что помогает в дальнейшем улучшать контент. GPS-навигация открывает новые возможности геолокации: аудиогиды могут предоставлять информацию, основанную на местоположении пользователя. Это особенно полезно для экскурсий по городам или музеям, где важно получить информацию о текущем местоположении. Также пользователи могут получать подсказки о том, как добраться до следующей точки маршрута. Системы синтеза речи позволяют аудиогидам генерировать речь на различных языках: пользователи могут получать информацию в формате аудио, а не просто читать текст на экране, что делает аудиогиды более доступными для широкой аудитории. Также системы распознавания речи позволяют осуществлять голосовое управление: пользователи могут задавать вопросы, используя голосовые команды, что делает взаимодействие более естественным и удобным.

Преимуществами интеграции аудиогидов с современными технологиями являются:

– персонализация: системы могут предлагать контент, основанный на интересах и местоположении пользователя;

– повышение вовлеченности: интерактивные элементы и персонализированный подход повышают уровень вовлеченности пользователей;

– улучшение доступности: аудиогиды особенно полезны для людей с ограниченными возможностями, обеспечивая равный доступ к информации.

Данный доклад посвящен аудиогиду по объектам Центрального ботанического сада НАН Беларуси и Электронному гиду по Национальной академии наук Беларуси, в которых применяются современные системы синтеза речи, основанные на алгоритмах искусственного интеллекта.

Интерактивные аудиогиды по научно-исследовательским центрам

Аудиогид «Электронный гид по Национальной академии наук Беларуси» разработан сотрудниками Объединенного института проблем информатики НАН Беларуси [2]. Он рассказывает о научных организациях, выдающихся ученых, интересных местах и научных открытиях, сделанных в Академии наук, раскрывая богатую историю и достижения белорусской науки.

Приложение позволяет пользователям погрузиться в историю Национальной академии наук Беларуси, начиная со дня ее основания и до наших дней. Оно включает информацию о ключевых событиях, выдающихся уче-

ных, значимых местах Академии наук, таких как исторические здания, лаборатории и музеи. Аудиогид структурирован по отраслям знаний, что облегчает поиск информации о конкретных направлениях науки.

Приложение доступно на трех языках: белорусском, русском и английском, что делает его доступным для широкой аудитории, включая людей с ограниченными возможностями зрения или слуха. Оно предоставляет полную информацию о каждой организации Академии наук, включая название, описание деятельности, специализацию, историю, ключевые достижения и контактные данные.

Аудиогид направлен на повышение интереса к науке, обеспечение общего доступа к информации и популяризацию Академии наук. Материал представлен в виде текстовых и озвученных описаний, а также фотоматериалов. Особенностью аудиогида является карта и система позиционирования по GPS, которая предоставляет интерактивную карту с указанием местоположения всех организаций Академии наук. Пользователи могут найти нужную организацию и увидеть ее местоположение на карте. Приложение также позволяет проложить маршрут для посещения организаций и интересных мест, связанных с наукой и самой Академией. Это удобно для организации экскурсий или научных мероприятий.

Электронная экскурсия по важным разработкам Академии наук представляет экспонаты, структурированные по отраслям и темам. Пользователи могут узнать о передовых исследованиях в различных областях науки – от фундаментальных до прикладных, от гуманитарных до технических – с помощью перекрестных ссылок на материалы. Дополнительной возможностью является голосовой поиск. Используя смартфон, пользователь может получить доступ к дополнительной информации об отдельном экспонате, включая фотографии и подробное описание, используя стандартные функции приложения или поиск по разработкам и объектам Национальной академии наук Беларуси. Аудиогид не только является интерактивным инструментом для передачи информации, но и открывает новые горизонты в понимании мира. Он позволяет ознакомиться с научной средой Беларуси, посетить Академию в более привлекательном формате и способствует популяризации науки, повышая интерес пользователей.

Цель разработки аудиогида для Центрального ботанического сада Национальной академии наук Беларуси – повышение удобства и свободы передвижения по местам Ботанического сада, быстрая навигация во время поиска и знакомство с наиболее интересными объектами на его территории, а также построение маршрута к выбранному объекту [3]. Центральный ботанический сад НАН Беларуси является самым крупным в стране центром по сохранению биоразнообразия живых растений, ведущим научным учреждением в области интродукции, акклиматизации, физиологии, молекуляр-

ной биологии, биотехнологии, биохимии и экологии растений. Созданный генофонд декоративных и хозяйственно-полезных интродуцированных растений (более 15 тыс. наименований) составляет национальное достояние Республики Беларусь. Площадь памятника природы республиканского значения составляет 93 га. Одно из старейших ботанических учреждений Беларуси имеет статус историко-культурной ценности 2-й категории, которая объединяет научно-исследовательскую деятельность в области сохранения биоразнообразия живых растений и прививает эстетическое отношение к природе Республики Беларусь.

Используя аудиогид, посетители Ботанического сада имеют возможность познакомиться с объектом садово-паркового искусства, сочетающим функции столичной достопримечательности и важнейшего культурно-просветительского, эколого-воспитательного и образовательного центра, а также дополнительной информацией о биоразнообразии живых растений и других объектах биологического значения с помощью телефона в простом и доступном формате.

Основные функции, реализованные в приложении:

– отображение территориальных зон Ботанического сада, экспозиций, секторов флор разных стран и непосредственно растений в текстовом и звуковом форматах;

– навигация: аудиогид отображает карту с объектами и прокладывает маршрут с помощью системы позиционирования по GPS;

– образовательный аспект: знакомство с новыми представителями флоры и экспонатами в удобной форме вызывает интерес у посетителей, желание узнать больше о том или ином объекте;

– возможность выбора языка (белорусский, русский, английский): позволяет туристам из разных стран понимать полученную информацию.

Созданный прототип экспериментальной автоматизированной системы информирования и навигации по объектам Центрального ботанического сада НАН Беларуси соответствует стандартам качественного приложения, основной целью которого является интерактивное представление объектов. Удобная навигационная система облегчает ориентацию на территории сада, а мобильные приложения и веб-версия аудиогида упрощают процесс посещения, делая его независимым от гидов.

Аудиогид имеет большое значение для повышения информационной доступности и навигационных возможностей в Центральном ботаническом саду НАН Беларуси. Разработка позволяет пользователям бесплатно знакомиться с экспонатами как на территории сада, так и виртуально, а организаторам – отслеживать наиболее популярные экспонаты и при необходимости обновлять их.

Системы синтеза речи как основной компонент аудиогидов

Этапы разработки аудиогидов включают следующие шаги:

1) планирование – определение целевой аудитории, выбор формата, языка интерфейса;

2) создание контента – подготовка текстовых описаний, аудио- и фотоматериалов, профессиональное озвучивание текстовых материалов;

3) разработка приложения – выбор платформы (Web, iOS, Android), разработка дизайна интерфейса и функционала;

4) реализация аудиогида – программирование макетного образца на различных версиях, тестирование работоспособности приложения на разных устройствах и платформах;

5) загрузка приложения в App Store и Google Play, размещение веб-версии на отдельном домене;

6) техническая поддержка приложения, обновление контента, анализ данных об использовании аудиогида для улучшения его качества.

Для реализации этапов были поставлены следующие задачи:

– разработка моделей хранения, изменения и выдачи электронных текстовых и речевых данных;

– создание прототипов экспериментальной автоматизированной системы для выдачи электронных мультимедийных данных об объектах и навигации по объектам Центрального ботанического сада НАН Беларуси и Национальной академии наук Беларуси;

– выявление и решение ресурсоемких задач синтеза речи по тексту для озвучивания текстовых описаний на первоначальном этапе реализации приложения.

Синтез речи в аудиогидах имеет ряд преимуществ в сравнении с привлечением профессиональных дикторов. Во-первых, он обеспечивает масштабируемость и скорость озвучивания: новые записи можно сгенерировать за считанные минуты без поиска диктора, правки в тексте моментально отражаются в готовых аудиофайлах, а параллельная обработка позволяет одновременно создавать версии для множества языков. Во-вторых, синтез облегчает многоязычность и локализацию – доступен выбор языков и акцентов, что упрощает озвучивание даже на редких языках. Третье – экономия затрат: синтез снижает затраты на студии, дикторов и переозвучку, что делает расходы предсказуемыми и легко оптимизируемыми. Четвертое – гибкость в настройке голоса и стиля: можно выбирать тембр, возраст, пол и эмоциональную окраску голоса, контролировать интонацию и паузы для лучшего восприятия на месте и применять динамическое озвучивание, адаптирующееся к контексту – от коротких подсказок до длинных исторических справок. Еще одним плюсом является надежность и предсказуе-

мость: качество остается стабильным без вариативности, свойственной живым записям (усталость, непреднамеренные оговорки), исчезают зависимости от графиков дикторов и студий. Наконец, с точки зрения юридических и операционных аспектов синтез упрощает вопросы прав и лицензирования – отпадает необходимость в длительных контрактах с дикторами – и позволяет быстро реагировать на внешние изменения, оперативно редактируя содержимое. Таким образом, синтез речи дает скорость, масштабируемость, экономию и гибкость при сохранении управляемого качества и легкой интеграции в цифровые сервисы аудиогидов.

Реализация системы синтеза речи в аудиогидах по Центральному ботаническому саду НАН Беларуси и НАН Беларуси

Системы синтеза речи на базе искусственного интеллекта (eng. artificial intelligence, AI) преобразуют текст в естественное звучание с помощью моделей машинного обучения. В их основе обычно лежат нейросетевые архитектуры, которые решают две ключевые задачи: преобразование текста в промежуточное акустическое представление и генерацию звуковой волны из этого представления. Такие системы обучаются на больших корпусах пар «текст – аудиозапись», что позволяет им усваивать закономерности произношения, интонации, ритма и тембра [4].

Современные AI-синтезаторы речи в основном построены на архитектурах глубокого обучения, таких как трансформеры и сверточные сети (CNN), которые вытеснили более ранние подходы, включая статистические параметрические модели и рекуррентные нейронные сети (RNN). Ключевые возможности таких систем включают управление просодией (интонацией, паузами, ударениями), регулировку скорости и громкости, а также добавление эмоциональной окраски. Многие решения поддерживают тонкую кастомизацию – пользовательские словари, правила транскрипции и корректировки произношения, а также мультилингвальность с учетом акцентов и региональных особенностей.

Генерация аудиосигнала из промежуточного представления выполняется либо нейросетевыми вокодерами (WaveNet-подобные модели, WaveGlow, HiFi-GAN и др.), либо end-to-end-моделями, которые сразу предсказывают форму звуковой волны. Современные вокодеры фокусируются на высокой естественности и низкой задержке, что делает возможным как пакетную генерацию, так и синтез в реальном времени. Развитие моделей и оптимизация вокодеров позволили достигнуть уровня, при котором синтезируемая речь становится все ближе к натуральной, а инструменты кастомизации дают возможность адаптировать голос под конкретные задачи.

Существуют также недостатки и ограничения AI-синтезаторов. Синтезированная речь может звучать менее естественно по сравнению с живым диктором, что снижает уровень вовлеченности пользователей. Системы могут неправильно произносить редкие слова, имена или термины, что приводит к недопониманию информации. Проблемы с просодией – неправильное управление интонацией, паузами и ударениями – влияют на восприятие и понимание текста. Для обучения моделей требуется значительное количество размеченных данных, что сложно обеспечить, особенно для менее распространенных и малоресурсных языков. Обучение и использование современных моделей синтеза речи требуют значительных вычислительных ресурсов, что является ограничивающим фактором для некоторых организаций. Кроме того, синтезаторы могут не учитывать региональные акценты и диалекты, что может снизить качество синтеза для определенных групп пользователей. Хотя современные модели имеют возможность добавлять некоторую эмоциональную окраску, они все еще не могут полностью передать нюансы человеческих эмоций, что может сделать аудиогид менее привлекательным. Следует учитывать и проблемы с доступом к технологиям или интернету, которые ограничивают использование аудиогидов, особенно в удаленных или сельских районах. Также для поддержания актуальности информации требуется регулярное обновление контента, что может быть трудоемким процессом.

В рамках разработки аудиогидов для синтеза речи на белорусском, русском и английском языках была выбрана архитектура VITS. Модель VITS (вариационный вывод с состязательным обучением для преобразования текста в речь) представляет собой одноступенчатую неавторегрессивную модель преобразования текста в речь, способную генерировать более естественно звучащую речь, чем существующие двухступенчатые модели, такие как Tacotron 2, Transformer TTS или даже Glow-TTS. Используя вариационный базис, VITS моделирует латентное пространство речевых признаков, отражая вариативность процесса генерации речи. Наличие состязательного обучения в VITS дополнительно улучшает процесс синтеза. Состязательное обучение включает в себя обучение сети-дискриминатора, способного различать реальную и синтезированную речь, в то время как сеть-генератор стремится генерировать речь, успешно обманывающую дискриминатор.

Архитектура VITS для систем синтеза речи обладает рядом значительных достоинств, которые делают ее особенно привлекательной для разработчиков. Модель генерирует более естественно звучащую речь по сравнению с традиционными моделями, что достигается благодаря использованию вариационного вывода и состязательного обучения. Это позволяет модели создавать звук, который ближе к человеческому произношению. Система имеет одноступенчатую неавторегрессивную архитектуру, что исключает

задержки, связанные с последовательной генерацией, и обеспечивает более быструю обработку. Эффективное моделирование латентного пространства речевых признаков позволяет учитывать вариативность и неопределенность в процессе генерации, что также способствует улучшению качества синтеза. Состязательное обучение, включающее наличие дискриминатора, значительно повышает качество синтезированной речи, так как модель стремится генерировать звук, который неотличим от реального. Гибкость и адаптивность VITS позволяют легко настраивать параметры синтеза, такие как интонация, скорость и эмоциональная окраска, что делает ее подходящей для различных приложений.

Модель также поддерживает мультилингвальность, что расширяет ее применение в многоязычных системах. Высокая производительность современных архитектур VITS обеспечивает быструю генерацию речи, что позволяет использовать их в реальном времени. Легкость интеграции с различными инструментами и библиотеками упрощает процесс разработки и развертывания. Кроме того, VITS демонстрирует хорошую устойчивость к шуму и артефактам, что улучшает качество синтезированной речи в сложных условиях. Поддержка кастомизации, включая возможность создания пользовательских словарей и правил транскрипции, улучшает точность произношения и адаптацию к специфическим требованиям. Все эти достоинства делают архитектуру VITS особенно привлекательной для разработки современных систем синтеза речи, обеспечивая высокое качество и гибкость в использовании.

Общая архитектура VITS состоит из апостериорного кодера, априорного кодера, декодера и стохастического предиктора длительности (рис. 1). Модули апостериорного кодера и дискриминатора декодера используются только во время обучения, а не для вывода.

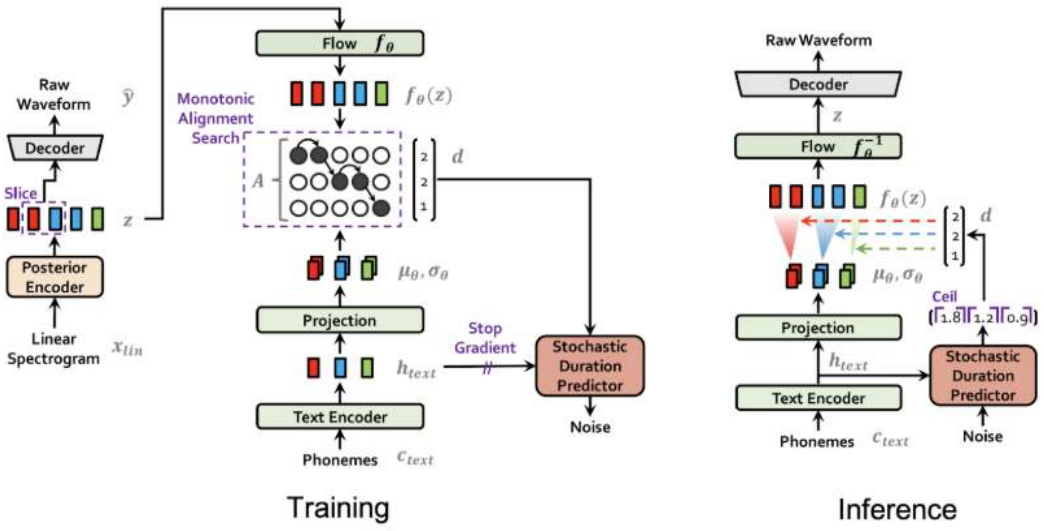


Рис. 1. Архитектура модели синтеза VITS

Для апостериорного кодера использованы 16 остаточных блоков WaveNet, состоящих из слоев расширенных сверток с управляемым активационным блоком и пропуском связи (рис. 2). Апостериорный кодер принимает на вход логарифмические спектрограммы амплитуд в линейном масштабе x_{lin} и генерирует скрытые переменные z с использованием 192 каналов. Основная задача апостериорного кодировщика – преобразование аудиоданных из пространства Mel-спектрограммы в пространство нормального распределения.

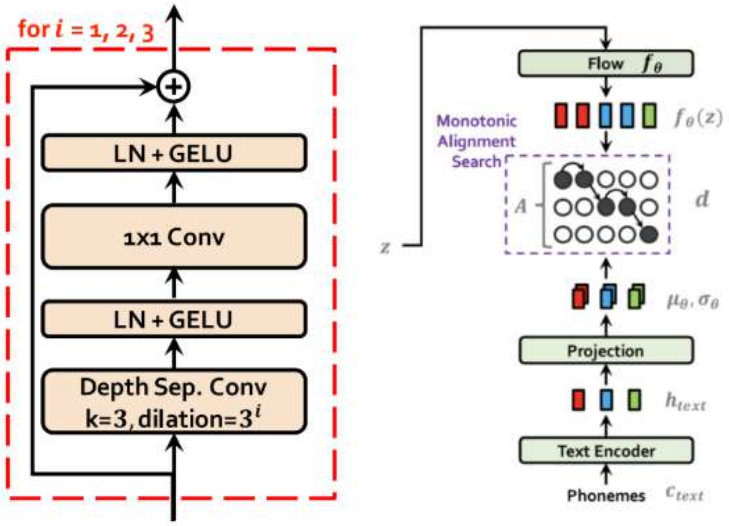


Рис. 2. Блок апостериорного кодера (слева), блок априорного кодера (справа)

Предварительный кодер состоит из текстового кодера, обрабатывающего входные фонемы c_{text}, и нормализующего потока f_θ, который повышает гибкость априорного распределения (рис. 2). Текстовый кодер представляет собой преобразователь, использующий относительное позиционное представление вместо абсолютного позиционного кодирования. Скрытое представление h_{text} из c_{text} можно получить через текстовый кодер и слой линейной проекции над текстовым кодером, который вычисляет среднее значение и дисперсию, используемые для построения априорного распределения. Нормализующий поток представляет собой стек аффинно-связанных слоев, состоящий из стека остаточных блоков WaveNet.

Декодер представляет собой генератор HiFi-GAN V1. Он состоит из стека транспонированных сверток, каждая из которых содержит модуль слияния мультирецепторных полей (MRF), как показано на рис. 3. Во время обучения на вход декодера подаются латентные переменные, генерируемые апостериорным кодером z, а во время вывода – латентные переменные, генерируемые априорным кодером $f\theta(z)$.

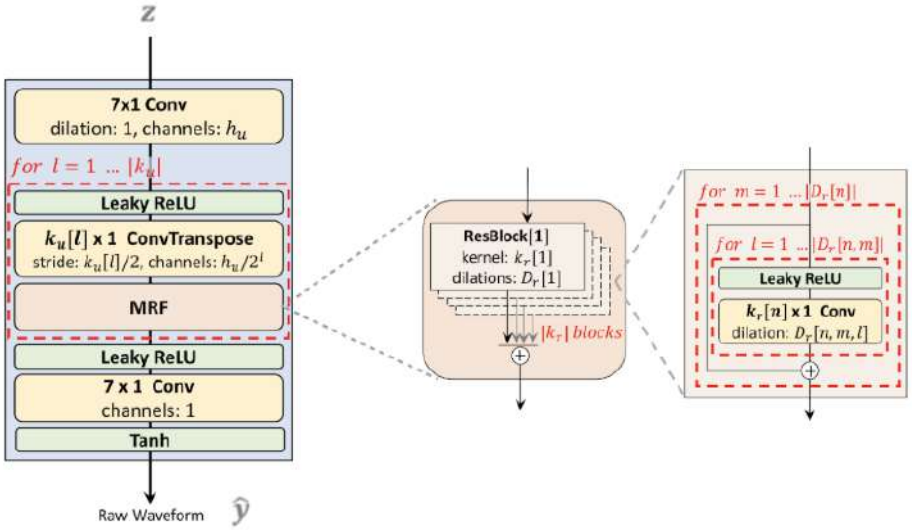


Рис. 3. Схема декодера архитектуры VITS

Стохастический предиктор длительности оценивает распределение длительности фонем по условному входному h_{text}. Для эффективной параметризации стохастического предиктора длительности накладываются остаточные блоки на расширенные и разделяемые по глубине сверточные слои (рис. 4). Нейронные сплайновые потоки, которые принимают форму обратимых нелинейных преобразований с использованием монотонных рационально-квадратичных сплайнов, также применяются в данном блоке к слоям связи. Нейронные сплайновые потоки улучшают выразительность преобразования при аналогичном количестве параметров по сравнению с обычно используемыми аффинными слоями связи.

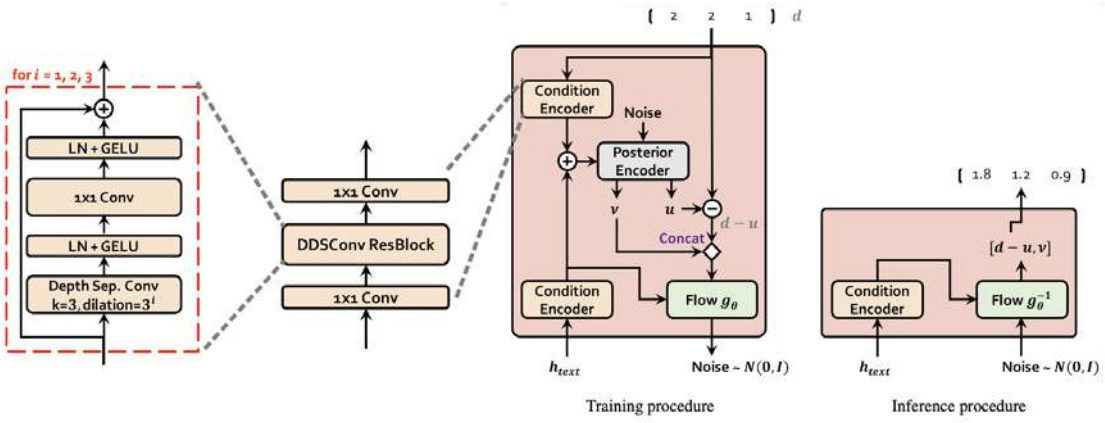


Рис. 4. Схема стохастического предиктора архитектуры VITS

Для озвучивания русскоязычных текстов в аудиогидах данные для обучения сети были взяты с датасета RUSLAN (Russian Spoken Language Corpus for Speech Synthesis). Корпус, состоящий из 22,200 высококачественных аудиозаписей речи с соответствующими текстовыми аннотация-

ми общей продолжительностью более 31 ч, является наибольшим по продолжительности датасетом для одного диктора на русском языке. Анализ отобранных данных проводился с целью получения статистической информации и понимания особенностей русского языка в контексте синтеза речи. Он включал оценку распределения фонетических единиц, длины фраз и других характеристик, которые могут влиять на качество синтезированной речи.

Для разработки системы синтеза речи на белорусском языке существенным шагом является сбор данных для обучения системы [5]. Учитывая ограниченную доступность специализированных датасетов для белорусского языка, в качестве основного источника был выбран датасет Mozilla CommonVoice. Он представляет собой большую коллекцию голосовых записей, собранных от добровольных участников, которые начитывают предложения на разных языках, включая белорусский. Датасет доступен для свободного использования и распространяется с открытой лицензией, что делает его ценным ресурсом для белорусскоязычного синтеза речи. Недостатком датасета от CommonVoice является тот факт, что он предназначен непосредственно для задач распознавания речи. В этом отношении записи в датасете могут быть непрофессиональными и содержать различные артефакты, шумы или недостатки, которые негативно влияют на качество синтеза речи.

В процессе сбора данных из CommonVoice была проведена предварительная фильтрация и отбор записей достаточной продолжительности и разнообразия различных речевых фонетических единиц, таких как фонемы, слова и фразы. На следующем шаге проведен анализ избранных данных для получения статистической информации и понимания особенностей белорусского языка в контексте синтеза речи. Анализ включал оценку распределения фонетических единиц, длительности фраз и других характеристик, которые могут оказать влияние на качество и естественность синтезируемой речи.

Модель VITS для обоих языков была обучена с помощью библиотеки Coqui TTS, популярного набора инструментов с открытым исходным кодом для синтеза текста в речь. Coqui TTS предоставляет полный набор инструментов и утилит для обучения и развертывания моделей TTS. Главные особенности библиотеки – генерация реалистичной и высококачественной речи, возможность настройки голосов для достижения желательного тембра, тональности и других параметров, поддержка различных форматов аудиовыхода, таких как WAV, MP3 и OGG, легкая интеграция в различные приложения и системы. В процессе обучения модель VITS использовала регистратор Weights and Biases. Это платформа для отслеживания и визуализации экспериментов машинного обучения. Она позволяет исследовате-

лям и разработчикам регистрировать и отслеживать ход обучения, метрики и производительность модели в режиме реального времени. Использование Coqui TTS и регистратора Weights and Biases способствовало эффективному проведению экспериментов, оптимизации модели и мониторинга производительности на протяжении всего процесса обучения. Для обучения был использован сервер с видеокартой Nvidia RTX4090. Оптимизация параметров происходила с помощью алгоритма AdamW – исправленной версии популярного алгоритма оптимизации Adam. Размер батча для обучения и оценки качества модели равнялся 74, время обучения модели составило 72 ч для каждого языка.

Таким образом, архитектура VITS, выбранная для разработки аудиогидов на белорусском, русском и английском языках, демонстрирует значительные преимущества в сравнении с традиционными моделями, обеспечивая более естественное звучание и высокую гибкость в настройках.

Заключение

Разработка аудиогидов для Центрального ботанического сада НАН Беларуси и НАН Беларуси представляет собой значительный шаг в популяризации науки и доступности информации для широкой аудитории. Использование современных систем синтеза речи на базе искусственного интеллекта, таких как архитектура VITS, позволяет создавать высококачественные аудиогиды, которые делают процесс обучения и ознакомления с научными достижениями более интерактивным и увлекательным.

Аудиогиды не только дают пользователям возможность глубже понять историю и достижения белорусской науки, но и способствуют улучшению навигации по объектам Ботанического сада и Национальной академии наук. Многоязычность приложений и возможность голосового поиска делают их доступными для различных групп пользователей, включая людей с ограниченными возможностями. Использование GPS и интерактивных карт значительно упрощает процесс посещения, позволяя пользователям самостоятельно планировать свой маршрут, а интеграция синтезаторов речи в аудиогиды позволяет обеспечить масштабируемость, гибкость и персонализацию контента. Правильный выбор архитектуры, тщательная подготовка текстов, внимание к качеству речи и пользовательским настройкам – ключевые факторы качественной системы.

Несмотря на существующие вызовы, такие как необходимость в больших объемах данных для обучения и возможные ошибки в произношении, применение технологий тонкого дообучения и постобработки помогает минимизировать эти недостатки. В результате аудиогиды становятся не только инструментом передачи информации, но и средством повы-

шения интереса к науке, что в конечном итоге способствует развитию научного мышления и образованности в обществе.

Список использованных источников

1. Аўдыягіды Krokam у індустрыі беларускага турызму / М. М. Слесарава, Ю. С. Гецэвіч, А. У. Унучак [і інш.] // Перспективы развития туризма в современных условиях: мировые тенденции и региональные контексты / Нац. агентство по туризму ; редкол.: И. Н. Воронович (предс.) [и др.]. – Минск : Колорград, 2023. – С. 454–463.

2. Лічбавая навігацыя ў свеце навукі: распрацоўка аўдыягіда для Нацыянальнай акадэміі навук Беларусі / М. М. Слесарава, А. А. Бакуновіч, Ю. А. Васілеўская, Ю. С. Гецэвіч // III Форум IT-Академграда «Искусственный интеллект в Беларуси» : докл., Минск, 10–11 окт. 2024 г. – Минск : ОИПИ НАН Беларуси, 2024. – С. 270–275.

3. Прататып аўтаматызаванай сістэмы выдачы інфармацыі і навігацыі па аб'ектах Цэнтральнага батанічнага сада НАН Беларусі / Д. І. Латышэвіч, М. А. Казлова, М. А. Слесарава [і інш.] // Информационные технологии в промышленности, логистике и социальной сфере (ITI*2023) : тез. докл. / Объед. ин-т проблем информатики НАН Беларуси ; под науч. ред. А. В. Тузикова, П. Н. Бибило, А. Г. Гривачевского. – Минск : ОИПИ НАН Беларуси, 2023. – С. 99–102.

4. Тэхналогіі аўтаматычнай апрацоўкі і аналізу маўлення з прымяненнем штучнага інтэлекту / Ю. С. Гецэвіч, В. В. Дыдо, Д. А. Бяляўскі [і інш.] // II Форум IT-Академграда «Искусственный интеллект в Беларуси» : докл., Минск, 12–13 окт. 2023 г. / НАН Беларуси; ОИПИ НАН Беларуси ; науч. ред. С. В. Кругликов, С. Н. Касанин. – Минск : ОИПИ НАН Беларуси, 2023. – С. 71–78.

5. Беларускамоўны сінтэз маўлення па тэксце: падыход на аснове глыбокага навучання / Я. С. Зяноўка, Д. А. Бяляўскі, Д. І. Латышэвіч [і інш.] // Развитие информатизации и государственной системы научно-технической информатизации (РИНТИ-2024) : докл. XXIII Междунар. науч.-техн. конф., Минск, 21 нояб. 2024 г. – Минск : ОИПИ НАН Беларуси, 2024. – С. 252–256.

УДК 378.6.091:004.8

Практика разработки ИИ-ассистента для коммуникационной системы высшего образования как элемент реализации государственной стратегии цифровой трансформации Республики Беларусь

Ю. А. Смышляева
Белорусский государственный технологический университет, Минск
E-mail: smyshlyeva.julia@outlook.com

Введение

Современный этап развития общества характеризуется интенсивным внедрением цифровых технологий во все сферы жизнедеятельности человека, включая систему образования. В условиях глобальной цифровизации особое значение приобретает интеграция технологий искусственного интеллекта (ИИ) в образовательный процесс, что позволяет обеспечить индивидуализацию обучения, повышение эффективности коммуникаций и оптимизацию управленческих решений в университетской среде.

Республика Беларусь, реализуя стратегические ориентиры устойчивого развития и технологической независимости, придает приоритетное значение цифровой трансформации национальной экономики и социальной сферы. Одним из ключевых направлений государственной цифровой политики выступает развитие интеллектуальных решений и платформ, направленных на повышение эффективности деятельности организаций и государственных институтов.

В рамках Государственной программы «Цифровое развитие Беларуси» на 2021–2025 годы предусмотрено поэтапное создание и внедрение технологий ИИ в различные секторы экономики, включая систему образования. Цифровая трансформация рассматривается в Республике Беларусь как один из ключевых факторов повышения конкурентоспособности национальной экономики и эффективности государственного управления. Стратегические ориентиры цифровой политики определены рядом базовых документов: Государственной программой «Цифровое развитие Беларуси» на 2021–2025 годы, Концепцией развития цифровой экономики и информационного общества до 2030 года, а также Национальной стратегией устойчивого развития.

Главной целью государственной цифровой политики является формирование единой цифровой экосистемы, обеспечивающей развитие технологий ИИ, больших данных, облачных решений и автоматизации бизнес-процессов. В сфере образования это выражается в необходимости перехода к модели «цифровой педагогики», внедрения интеллектуальных платформ и инструментов, способствующих эффективному взаимодействию всех участников образовательного процесса.

Одним из практических примеров реализации данных стратегических задач выступает проект разработки ИИ-ассистента для высшего учебного заведения – УО «Белорусский государственный технологический университет».

Реализация проекта отражает приоритеты национальной стратегии цифровой трансформации, направленной на создание инновационной образовательной среды, повышение качества сервиса и развитие компетенций цифрового общества в Республике Беларусь [1].

Цифровая трансформация выступает одним из ключевых направлений государственной политики Республики Беларусь, направленной на повышение конкурентоспособности национальной экономики и эффективности системы государственного управления. Согласно Указу Президента Республики Беларусь от 7 мая 2020 г. № 156 «О приоритетных направлениях научной, научно-технической и инновационной деятельности на 2021–2025 годы» развитие цифровых информационно-коммуникационных и междисциплинарных технологий, включая ИИ, большие данные, робототехнику и технологии «умного» города, признано приоритетным направлением государственной инновационной политики.

В соответствии с Государственной программой «Цифровое развитие Беларуси» на 2021–2025 годы система образования включена в число ключевых секторов цифровой модернизации. Министерством образования Республики Беларусь определен курс на создание «умного университета» – модели образовательного учреждения, основанной на интеграции интеллектуальных систем, автоматизации административных и учебных процессов, а также персонализации образовательных траекторий обучающихся.

В Указе Президента Республики Беларусь № 156, Концепции развития цифровой экономики и информационного общества до 2030 года, а также Государственной программе «Цифровое развитие Беларуси» на 2021–2025 годы обозначены ключевые направления и задачи цифровизации сферы образования. Среди них можно выделить следующие приоритеты:

– развитие цифровой инфраструктуры образовательных учреждений, включающее создание и модернизацию национальных платформ дистанционного обучения, интеграцию интеллектуальных систем управления

образовательным процессом и обеспечение технологической совместимости различных цифровых сервисов;

– внедрение технологий ИИ, больших данных и облачных решений для повышения эффективности администрирования, автоматизации бизнес-процессов и совершенствования аналитической поддержки управленческих решений;

– формирование цифровой педагогики, основанной на применении инновационных методов обучения и адаптивных технологий, способствующих персонализации образовательных траекторий и развитию критического мышления обучающихся;

– повышение уровня цифровых компетенций студентов, преподавателей и административного персонала, формирование культуры работы в цифровой среде и развитие навыков цифровой безопасности;

– укрепление взаимодействия науки, образования и бизнеса в сфере разработки и внедрения ИИ-решений, стимулирование инновационной активности университетов и трансфера технологий;

– обеспечение информационной безопасности и защиты персональных данных в рамках функционирования цифровой образовательной экосистемы.

ИИ рассматривается как базовая технология цифровой трансформации, обеспечивающая интеллектуализацию процессов управления и коммуникации. Его внедрение в систему образования позволяет:

– автоматизировать взаимодействие со студентами и абитуриентами;

– персонализировать образовательные программы;

– повысить качество и скорость принятия управленческих решений.

Современные тенденции цифровизации образования, усиление конкуренции среди высших учебных заведений и рост ожиданий со стороны абитуриентов формируют необходимость внедрения инновационных инструментов маркетинговых коммуникаций. В статье представлена практика разработки ИИ-ассистента для коммуникационной системы УО «Белорусский государственный технологический университет». Разработанный цифровой инструмент обеспечивает круглосуточную консультационную поддержку абитуриентов, персонализированные рекомендации, мультиязычную навигацию и автоматизацию процессов приемной кампании.

Понятие и сущность ИИ-ассистентов

Интеллектуальные ассистенты, базирующиеся на технологиях ИИ, представляют собой программные системы, способные выполнять задачи, требующие обработки данных, распознавания естественного языка, принятия решений и адаптивного обучения. В контексте высшего образования они выступают инструментами поддержки административных процессов,

информационного сопровождения абитуриентов и студентов, а также персонализации образовательного опыта. Основными характеристиками ИИ-ассистентов являются способность к адаптивному взаимодействию, обработка больших объемов информации, интеграция с существующими цифровыми платформами университета и возможность самообучения на основе накопленных данных. Теоретически применение таких систем основано на моделях когнитивного и машинного обучения, что обеспечивает их функциональную гибкость и потенциал для оптимизации образовательных коммуникаций.

Международный опыт внедрения ИИ в образовательные процессы

Применение ИИ в образовательной среде демонстрирует значительный прогресс в ведущих университетах и на платформах EdTech. Например, Массачусетский технологический институт (MIT) и Стэнфордский университет внедряют интеллектуальных чат-ботов для поддержки абитуриентов и студентов, автоматизации административных запросов и сопровождения учебного процесса. EdTech-платформы, такие как Coursera и Udemy, используют алгоритмы ИИ для адаптивного тестирования, прогнозирования успеваемости и персонализированных рекомендаций курсов. Анализ международного опыта показывает, что интеграция ИИ позволяет повышать оперативность предоставления информации, увеличивать вовлеченность пользователей, снижать нагрузку на сотрудников и создавать условия для более качественного и индивидуального образовательного процесса [2].

Современная система высшего образования переживает этап активной цифровизации, что создает новые возможности для оптимизации учебного процесса, повышения качества коммуникаций и эффективности управленческих решений. Одним из ключевых инструментов цифровой трансформации университетов становятся технологии ИИ. Их внедрение открывает перспективу автоматизации рутинных задач, ускорения обработки запросов студентов и абитуриентов, а также формирования персонализированных образовательных траекторий.

ИИ-ассистенты позволяют значительно снизить административную нагрузку на сотрудников, обеспечивая круглосуточную поддержку и более точное распределение запросов по категориям сложности. Алгоритмы машинного обучения дают возможность анализировать большие массивы данных, выявлять закономерности в поведении обучающихся и принимать обоснованные управленческие решения. В результате взаимодействие студентов с образовательной средой становится более гибким, индивидуализированным и информативным, а университет получает новые возможно-

сти для стратегического планирования и повышения качества образовательных услуг.

Вместе с тем использование ИИ-технологий сопряжено с рядом ограничений и потенциальных рисков. Алгоритмы могут допускать ошибки или искажения данных, что способно приводить к некорректным выводам и решению проблем. Некоторые универсальные системы часто недостаточно адаптированы к культурным и языковым особенностям пользователей, что снижает эффективность коммуникаций и персонализации. Кроме того, активное использование цифровых технологий требует повышенного внимания к вопросам информационной безопасности и защиты персональных данных. Также существует риск технологической зависимости университетов от внешних поставщиков инфраструктуры и программного обеспечения.

Интеграция ИИ в высшее образование представляет собой баланс между очевидными преимуществами и необходимостью системного управления рисками. Эффективное использование ИИ возможно лишь при условии продуманного внедрения, прозрачности алгоритмов, соблюдения этических норм и постоянного контроля качества работы интеллектуальных систем. ИИ не заменяет человеческое участие, а усиливает организационный и интеллектуальный потенциал образовательного учреждения, создавая условия для формирования инновационной и адаптивной цифровой образовательной среды.

Практика разработки ИИ-ассистента для высшего образования

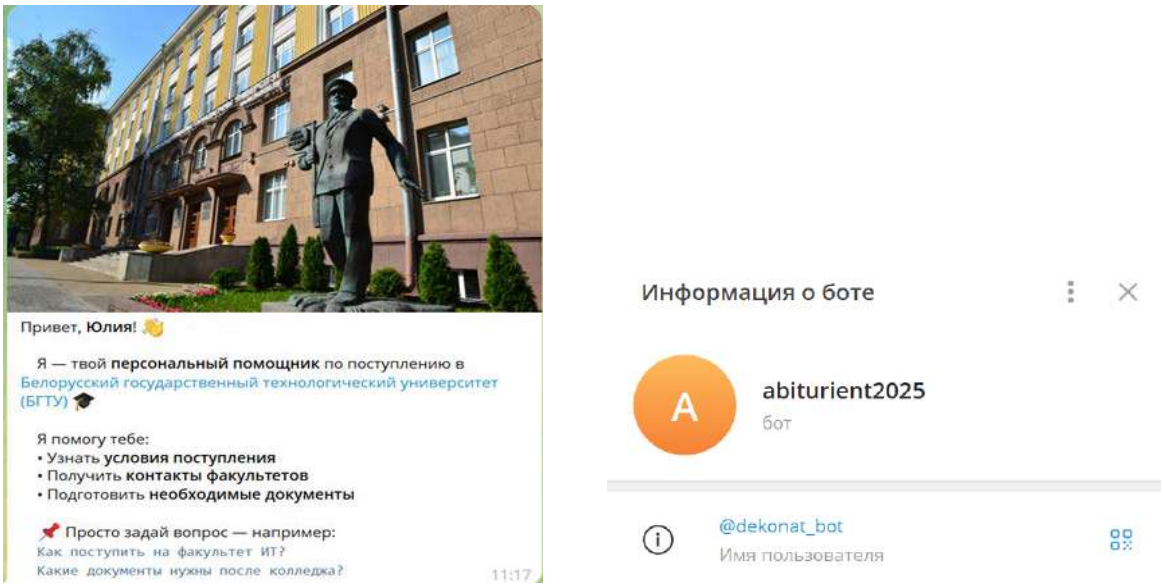
Инициатива по созданию проекта обусловлена объективной необходимостью повышения эффективности и качества коммуникаций между образовательной организацией и потенциальными абитуриентами в условиях активной цифровизации системы высшего образования. Увеличение количества обращений, разнообразие форм запросов и рост требований к оперативности предоставления информации выявили ограниченность традиционных каналов взаимодействия (телефонных линий, электронной почты, очных консультаций). В данных условиях возникла потребность в разработке интеллектуального решения, способного обеспечить автоматизацию обработки запросов, персонализацию коммуникаций и круглосуточную доступность информационной поддержки. Реализация проекта ИИ-ассистента стала ответом на стратегический запрос по оптимизации внутренних процессов и повышению уровня цифровой зрелости университета.

Проведен анализ потребностей абитуриентов и внешних пользователей коммуникационной системы университета. При разработке использованы технологии обработки естественного языка (NLP), мультиязычные интерфейсы, персонализированные алгоритмы рекомендаций без интегра-

ции с CRM-системой, официальным сайтом и мессенджерами. Архитектура решения предусматривает модульность и возможность последующего масштабирования [3].

Оценка эффективности на текущем этапе основана на анализе предварительных данных и динамических показателей, характеризующих активность пользователей, количество взаимодействий и структуру обращений. Дальнейшее развитие исследования предполагает накопление эмпирических данных и проведение сравнительного анализа фактических и прогнозных результатов, что позволит уточнить количественные показатели эффективности и подтвердить выдвинутые гипотезы.

Разработанный для коммуникационной системы УО «Белорусский государственный технологический университет» ИИ-ассистент «abiturient2025» позволит оптимизировать процессы взаимодействия с абитуриентами и повысить эффективность работы приемной комиссии (рисунок). Экспериментальная апробация продемонстрировала оперативность ответов ИИ-ассистента за счет автоматизации ответов на типовые запросы и распределения обращений по уровням сложности. Одновременно наблюдалось повышение скорости и качества консультаций, что выражалось в увеличении доли полностью удовлетворенных пользователей и сокращении среднего времени отклика.



Пример страницы ИИ-ассистента

Сравнительный анализ показал, что по функциональной зрелости и интеграционным возможностям разработанное решение сопоставимо с ИИ-платформами, применяемыми в ведущих российских университетах, при этом превосходит типовые конструкторы, используемые в большин-

стве белорусских вузов, по глубине персонализации, мультиязычности и аналитическому потенциалу.

Количественные результаты пока носят характер динамических наблюдений, полученные данные свидетельствуют о высокой адаптивности ИИ-ассистента и перспективности его дальнейшего масштабирования. В дальнейшем планируется проведение углубленного анализа метрик эффективности и оценка влияния внедренной системы на качество коммуникаций вуза в долгосрочной перспективе. Таким образом, результаты исследования подтверждают высокую адаптивность и эффективность внедренного решения, а также демонстрируют перспективность его тиражирования в образовательной среде Республики Беларуси [4].

Архитектура, технологические решения и функциональные возможности ИИ-ассистента

Архитектура ИИ-ассистента построена по модульному принципу, обеспечивая гибкость, масштабируемость и возможность интеграции с существующей цифровой инфраструктурой университета. Система включает несколько ключевых компонентов, каждый из которых отвечает за определенные аспекты взаимодействия с пользователями и обработки образовательных данных.

Модель обработки естественного языка (NLP) реализована с использованием современных трансформеров (BERT, GPT), что обеспечивает точное распознавание, классификацию и генерацию ответов на запросы пользователей. Данный компонент поддерживает многоязычное диалоговое взаимодействие и адаптируется к различным сценариям коммуникации с абитуриентами, студентами и сотрудниками университета.

Алгоритмы машинного обучения применяются для прогнозирования образовательных потребностей, персонализации рекомендаций и оптимизации качества взаимодействия. Используются методы supervised и reinforcement learning, что позволяет непрерывно улучшать точность ответов и адаптировать сценарии взаимодействия под индивидуальные особенности пользователей.

Интеграция реализована посредством REST API, что позволяет синхронизировать данные и обеспечивать оперативный обмен информацией между системами. Система хранения данных включает реляционную базу PostgreSQL для структурированных данных пользователей и обращений. Такая комбинация обеспечивает надежность, масштабируемость и возможность аналитической обработки больших массивов данных.

Интерфейс ИИ-ассистента представлен мультиязычными веб- и мобильными приложениями, разработанными с использованием Node.js. Это гарантирует удобство и доступность взаимодействия для всех категорий

пользователей. Функциональные возможности системы охватывают несколько ключевых направлений:

1. ИИ-ассистент сопровождает абитуриентов и студентов, предоставляя справочную информацию, автоматизируя ответы на типовые запросы, сопровождая процесс поступления и поддерживая адаптацию новых студентов. Кроме того, система поддерживает учебный процесс, включая напоминания о сроках сдачи заданий, консультации по расписанию и учебным материалам, помощь в выборе курсов и образовательных модулей.

2. Аналитический модуль системы собирает и обрабатывает данные о поведении пользователей, формируя индивидуальные рекомендации и прогнозы академической успеваемости, что способствует персонализации образовательных траекторий и поддержке принятия решений как студентами, так и администрацией.

3. Предложенная архитектура и функциональная реализация ИИ-ассистента создают интегрированную интеллектуальную среду, способствующую оптимизации образовательных процессов, повышению эффективности работы университета и персонализации взаимодействия с пользователями.

Заключение

Внедрение ИИ-ассистента в систему высшего образования представляет собой значимый шаг в направлении цифровой трансформации образовательных процессов. Использование технологий ИИ способствует оптимизации коммуникации между университетом, абитуриентами и студентами, повышает эффективность административного управления и качество предоставляемых образовательных услуг [5].

Интеллектуальные системы на основе ИИ обеспечивают персонализированное сопровождение обучающихся, расширяют возможности анализа данных и повышают адаптивность образовательной среды. Вместе с тем успешная реализация подобных проектов требует комплексного подхода, включающего разработку этических стандартов, обеспечение защиты персональных данных и формирование цифровых компетенций у всех участников образовательного процесса. Такая технология является не только технологическим, но и стратегическим направлением, определяющим конкурентоспособность и инновационный потенциал образовательных учреждений в условиях цифровой экономики.

Список использованных источников

1. О Государственной программе «Цифровое развитие Беларуси» на 2021–2025 гг. : постановление Совета Министров Респ. Беларусь от 2 февр. 2021 г. № 66 // Национальный правовой Интернет-портал Республики Беларусь. – URL: https://pravo.by/document/?guid=3871&p0=C22100066 (дата обращения: 15.10.2025).

2. Бессмертный, И. А. Интеллектуальные системы : учебник и практикум для вузов / И. А. Бессмертный, А. Б. Нугуманова, А. В. Платонов. – М. : Юрайт, 2024 – 243 с.

3. Искусственный интеллект в Беларуси: какие технологии разрабатывают и как их будут регулировать // Национальный правовой Интернет-портал Республики Беларусь. – URL: https://pravo.by/novosti/obshchestvenno-politicheskie-i-v-oblasti-prava/2025/june/88987/ (дата обращения: 15.10.2025).

4. Government Artificial Intelligence Readiness Index 2022 // Oxford Insights. – 21 January 2023. – URL: https://static1.squarespace.com/static/58b2e92c1e5b6c828058484e/t/639b495cc6b59c620c3ecde5/1671121299433/Government_AI_Readiness_2022_FV.pdf (дата обращения: 10.09.2025).

5. Посещение выставки научных достижений «Беларусь интеллектуальная» // Офиц. Интернет-портал Президента Республики Беларусь. – URL: https://president.gov.by/ru/events/poseshchenie-vystavki-nauchnyh-ostizheniy-belarus-intellektualnaya (дата обращения: 10.09.2025).

УДК 004.8:35.08

Использование генеративного искусственного интеллекта в практике работы государственных органов и организаций

В. А. Степанов
НИИ теории и практики государственного управления
Академии управления при Президенте Республики Беларусь,
Минск

Введение

В белорусском обществе растет интерес к технологиям искусственного интеллекта (ИИ). Контент-анализ публикаций информагентства БелТА показывает, что если в январе–июне 2024 г. на сайте вышло 22 материала с тэгом «ИИ», то за аналогичный период 2025 г. – уже 75. При этом большинство публикаций (около 40 %) посвящены теме внедрения искусственного интеллекта. Это косвенно свидетельствует о том, что и государство, и общество осознали практический потенциал технологий, которые, как утверждают аналитики агентства McKinsey, могут сыграть для социума такую же роль, какую паровой двигатель сыграл во время промышленной революции XIX в. [1].

Но насколько активно и, главное, результативно инструменты ИИ внедряются в Беларуси? Чтобы ответить на этот вопрос, были изучены повседневные профессиональные практики сотрудников государственных органов и организаций. Этот сегмент был выбран для анализа по ряду причин. Во-первых, он хорошо поддается оптимизации за счет внедрения ИИ: в среднем госслужащий тратит на выполнение базовых административных задач (бюрократические процедуры) до 30 % рабочего времени [2]. Во-вторых, это консервативная, инерционная система, в которой инновации внедряются, как правило, централизованно, когда накоплен достаточный опыт, приняты соответствующие правовые основы, установлены технические стандарты. Такая конфигурация ожидаемо порождает неформальные, параллельные социальные практики, когда технология осваивается в обход институциональных норм.

Методы исследования и результаты

Используя анонимное электронное анкетирование [3], был сделан опрос в сентябре 2025 г. 53 респондентов – руководителей и специалистов

республиканских и местных органов государственного управления (около 40 % от общего числа), редакций государственных СМИ (около 20 %), иных организаций и предприятий (почти 40 %). Географически были представлены все регионы с небольшим перевесом респондентов из Минска (30 %). Основная возрастная группа – 30–50 лет (порядка 80 %).

Опрос проводился в основном среди слушателей программ переподготовки и повышения квалификации АУпПРБ. Таким образом, несмотря на относительно небольшой размер выборки, исследование охватило преимущественно активную, заинтересованную в инновациях часть госслужащих и сотрудников госорганизаций.

При формулировании вопросов ориентировались на исследования [4–7], что позволило соотнести полученные результаты с наработками коллег и выявило следующее:

1. За последний год ИИ-технологии в той или иной степени повлияли на работу большинства респондентов (74 %) вне зависимости от того, использовали ли они их сами, или нет. Каждый десятый сообщил, что влияние очень сильное.

2. Говоря о позитивных эффектах, участники опроса отмечали, что благодаря ИИ они стали выполнять работу быстрее (72 %), появились новые форматы работы (68 %), повысилась их креативность (64 %). В то же время освоение ИИ-инструментов требует времени и дополнительных трудозатрат (34 %). Респонденты утверждают, что у них стало уходить больше усилий на перепроверку информации и редактирование текстов (25 %).

3. Большинство респондентов (76 %) используют в своей профессиональной деятельности общедоступные сервисы ИИ; более 40 % делают это регулярно – ежедневно или еженедельно. Хотя лишь один участник опроса упомянул о прямом запрете на применение ИИ в организации, правила информационной безопасности, такие как конфиденциальность обрабатываемых сведений, запрет на VPN и др., являются серьезным препятствием, которое не дает специалистам обращаться к помощи ИИ-сервисов на работе. Другие значимые причины – не хватает знаний и компетенций (21 %), нет возможности оплачивать приложения (19 %).

4. Для решения рабочих задач в основном используются универсальные чат-боты (74 %). Далее следуют специализированные инструменты для создания графики (25 %), работы с текстом (19 %). Самые популярные сервисы, которые указали респонденты, – чат-боты ChatGPT и Deepseek. Помимо них упомянуты YandexGPT, Gigachat, Copliot, Grok, Gemini, Perplexity, Qwen. Из генераторов изображений чаще указывали «Кандинский», «Шедеврум» и Leonardo. Системы для оптимизации рабочих процессов (подготовки протоколов, создания презентаций и др.) используются

реже (9 %). Несколько респондентов сообщили о применении корпоративных ИИ-систем, разработанных для нужд организации.



Рис. 1. Используемые ИИ-сервисы (% числа опрошенных)

5. Сотрудники госорганов в большинстве своем используют общедоступные бесплатные сервисы (72 %). Небольшая доля опрошенных (13 %) ориентируется на бесплатные приложения, но выборочно оплачивает самые нужные инструменты. Прецеденты, когда участники опроса оформляли регулярную подписку сами или при помощи работодателя, который компенсировал издержки, также есть, но они единичны.

6. Если говорить о способах взаимодействия с ИИ-инструментами, то нельзя однозначно выделить предпочтительный интерфейс. Часть респондентов заходит на сайт (40 %), часть – использует Telegram-боты (39 %), часть – фирменное приложение (32 %). Можно лишь констатировать, что практически никто не работает с ИИ как с технологией, «вшитой» в экосистему приложений (Microsoft, Google).

7. Чаще ИИ используют как «поисковик» и «редактор» (инструмент для работы с текстом). Так, среди задач, для решения которых нейросети применяются регулярно, отмечены следующие: поиск необходимой для работы информации, сбор материала по теме, создание черновиков документов, подготовка проектов речей, редактирование и перевод текстов. Точечно респонденты используют ИИ в роли «креативщика» (генерирование идей), «аналитика» (обработка данных, реферирование текстов) и «дизайнера» (создание графики, презентаций). Единицы создают с помощью ИИ видео, эксплуатируют его как эксперта-консультанта и пытаются автоматизировать бизнес-процессы, например работу с обращениями граждан.

8. В большинстве ведомств и организаций нет правил, регламентирующих использование ИИ, а также не планируется их создание (60 %). Тем не менее ряд участников опроса сообщили о подготовке таких документов (26 %). В некоторых случаях существуют устные распоряжения руководства, не оформленные в виде документа (11 % опрошенных).

9. Говоря о сложившихся социальных практиках, участники опроса отмечали, что вынуждены перепроверять факты и идеи, которые им предлагает ИИ (66 %). Хотя в целом тратят меньше времени и усилий на выполнение задач, зная, что могут воспользоваться для их решения ИИ-сервисами (64 %). У этих процессов есть и обратная сторона: от 15 до 20 % респондентов сказали, что им случалось делегировать выполнение задачи ИИ, вместо того чтобы взаимодействовать с коллегами, самим вникать в проблему или осваивать нужный навык. Любопытно, что 28 % респондентов признались, что им случалось представлять созданный ИИ контент как свой собственный. Но практика этичного использования нейросетей все же распространеннее: почти 60 % сообщали, что выполняли задачи с помощью ИИ.

10. Респонденты не боятся, что развитие ИИ-технологий приведет к гибели их профессии. Абсолютное большинство (около 80 %) признают, что изменится каждодневная практика выполнения работы, ее формат, но убеждены, что принципиальных изменений не произойдет.

11. В исследуемой сфере пока не сложился консенсус относительно технологической базы для ИИ, который будет внедряться в государственных органах и организациях. Так, 43 % считают, что Беларусь будет локализовать разработки дружественных стран (Китая и России); 26 % полагают, что будут разработаны собственные LLM и чат-боты, которые будут рекомендованы к использованию в госсфере; наконец, 21 % за то, чтобы наша республика использовала открытые (open-source) технические решения, которые будут доработаны под местный контекст, речь идет в основном о дообучении моделей, fine-tuning.

12. Как респонденты видят шаги по внедрению технологий? В приоритетах (по мере убывания количества голосов): формирование у сотрудников ИИ-грамотности за счет программ обучения; принятие документа, который регламентировал бы корректное и этичное использование ИИ; выделение бюджета для оплаты проприетарных версий ИИ-приложений; выделение бюджета на разработку и внедрение корпоративной ИИ-системы. Типичным бюрократическим процедурам (назначение ответственных лиц, пересмотр должностных инструкций, принятие стратегий и др.) отдан низкий приоритет или заявлено о том, что в них нет необходимости.



Рис. 2. Приоритетность шагов по внедрению ИИ

13. Если бы у респондентов была возможность пройти обучение по теме использования ИИ, то наибольший практический интерес у них вызвали вопросы аналитики и работы с данными (64 %), создания и редактирования текстов с помощью ИИ (60 %), верификации ИИ-контента (55 %), информационной безопасности (51%), создания и обработки изображений (49 %), видео и аудио (45 %). Темы, связанные с различными аспектами внедрения ИИ в работу организации (техническими, юридическими, организационными), интересуют лишь около трети участников опроса.

14. Говоря о рисках, актуальных при внедрении ИИ в работу организации, респонденты делают акцент на вопросах кибербезопасности (74 % отметили риск как важный), информационно-психологической безопасности (66 %), профессиональной деградации и несамостоятельности сотрудников, которые делегируют решение задач ИИ (57 %); на риске системного сбоя, коллапса в работе организации из-за делегирования выполнения ключевых задач ИИ (51 %); на политических коллизиях, вызванных давлением со стороны владельцев ноу-хау и др. (43 %). Интересно, какие риски признаны самыми неважными: снижение количества и качества межличностного общения вследствие внедрения ИИ, а также негативное воздействие на окружающую среду.

Обсуждение

Как видно из ответов, белорусские государственные органы и организации используют ИИ для решения профессиональных задач, применяя в основном общедоступные чат-боты (более 70 % опрошенных). Можно сказать, что организации находятся на этапе стихийного массового освоения этих технологий.

Аналогичную картину фиксируют и зарубежные исследования. Так, согласно опросу университета Мельбурна и компании KPMG, охватившему более 48 тыс. специалистов из 47 стран [4], почти 60 % респондентов регулярно используют ИИ в работе, причем треть делает это ежедневно. Сотрудники не ждут, пока менеджеры разработают стратегии внедрения ИИ, определят этические рамки или обеспечат их корпоративными системами, а действуют самостоятельно, используя общедоступные, преимущественно бесплатные сервисы. Около 70 % опрошенных применяют чат-боты (ChatGPT, Copilot, Claude). Далее следуют голосовые помощники (Siri, Alexa), затем генераторы изображений, видео и аудио (DALL-E, Canva).



Рис. 3. Оценка долгосрочных рисков, связанных с внедрением ИИ

Как было показано выше, для белорусских госорганизаций ИИ – это прежде всего умный поисковик, точнее, референт, готовящий справки, и редактор, способный написать и обработать текст по заданию. Такая фиксация на тексте понятна – большая часть процессов в исследуемых организациях связана с подготовкой документов. К тому же чат-боты, работающие с текстом, отличаются низким технологическим и финансовым «порогом вхождения» для пользователей. Такая картина коррелирует с международной практикой. Например, опрос сотрудников бюджетных организаций и местных органов власти США, проведенный О. Гизеке в 2024 г. [5], показал, что 8 из 10 госслужащих, которые используют ИИ, применяют чат-боты, а около 40 % – приложения для написания и редактирования текстов (writing assistants).

Из белорусских исследований, в которых поднимается схожая проблематика, стоит отметить замеры агентства MASMI. Опрос 50 маркетологов из Минска и 500 горожан по всей стране в апреле–июне 2025 г. [8] показал, что почти 80 % специалистов используют ИИ в повседневной деятельности. С помощью нейросетей они создают тексты (81 %), графику (69 %), ищут информацию (69 %), делают прогнозы (33 %), управляют соцсетями (31 %) и др. Ключевые инструменты – ChatGPT (92 %) и MidJourney (53 %). Как видим, если говорить о задачах, которые выполняются при помощи ИИ, то госструктуры и бизнес идут в одном фарватере.

Опрос госслужащих в ОАЭ [7] подсветил закономерность: специалисты, слабо знакомые с ИИ, используют чат-боты для решения простых задач, таких как создание контента или подготовка писем. Искушенные пользователи применяют нейросети для анализа данных, при планировании и управлении рисками. Если смотреть на компетенции, которые бы хотели освоить белорусские респонденты, то можно сказать, что после этапа «знакомства» (2022–2024 гг.) организации уже созрели для этапа «продвинутого использования».

Анкетирование демонстрирует прагматизм сотрудников госорганизаций: они признают, что ИИ изменит их каждодневные практики, скорректирует процедуры, но не боятся, что он уничтожит профессию. Многие признают, что стали больше полагаться на нейросети и прилагать меньше усилий. Есть понимание связанных рисков: в лучшем случае – профессиональная деградация, в худшем – организационный коллапс, если ИИ ошибется при принятии ключевых решений. Но пока все это ощущается как весьма отдаленная перспектива.

Зарубежный опыт показывает, что ИИ-инструменты способны повысить эффективность работы сотрудников госорганизаций. Около половины респондентов О. Гизеке (43,6 %) сообщили о росте продуктивности. Белорусы говорят о том же: растет скорость выполнения работы, появляются новые форматы, есть ресурсы для выполнения креативных задач. Следующим этап должно стать институциональное внедрение, это когда не просто увеличивается продуктивность конкретного работника, а меняются бизнес-процессы в масштабах организации. Обзор мировой практики использования ИИ в госорганах демонстрирует, что он может применяться либо для отладки внутренних процессов (бэк-офис), либо при непосредственном соприкосновении с гражданами, потребителями услуг (фронт-офис). Пример системы первого типа – «закрытый аналог ChatGPT» – сингапурский бот Pair, которым пользуется порядка 100 ведомств. Пример использования ИИ для фронт-офиса – британский бот GOV.UK Chat, который консультирует пользователей по вопросам открытия бизнеса и налогообложению.

На этапе индивидуального массового освоения ИИ руководителей больше беспокоят вопросы обучения персонала ИИ-грамотности и шлифовки практики этичного использования нейросетей. Но риски в области информационной безопасности, вполне отчетливо осознаваемые, подталкивают менеджмент к разработке и внедрению собственных корпоративных ИИ-систем, которые обучены на данных организации. Несмотря на наличие прецедентов, когда отдельные государства вступали в ИИ-гонку и разрабатывали собственные LLM (Egune в Монголии, Apertus в Швейцарии и др.), большинство респондентов склоняются к тому, что белорусский путь – это опора на технологическую мощь стран-союзниц (Китая, России) и доработка их ИИ-сервисов с учетом локальной специфики.

Список использованных источников

1. Superagency in the workplace: Empowering people to unlock AI's full potential. – URL: https://www.mckinsey.com/capabilities/mckinsey-digital/our-insights/superagency-in-the-workplace-empowering-people-to-unlock-ais-full-potential-at-work (date of access: 29.09.2025).

2. How much time and money can AI save government? – URL: https://www2.deloitte.com/content/dam/insights/us/articles/3834_How-much-time-and-money-can-AI-save-government/DUP_How-much-time-and-money-can-AI-save-government.pdf (date of access: 29.09.2025).

3. Онлайн-опрос «Использование генеративного ИИ в практике работы государственных органов и организаций» : [сайт]. – URL: https://vk.cc/cPkZpd (дата обращения: 29.09.2025).

4. Trust, attitudes and use of artificial intelligence. A global study 2025. – URL: https://assets.kpmg.com/content/dam/kpmgsites/xx/pdf/2025/05/trust-attitudes-and-use-of-ai-global-report.pdf (date of access: 29.09.2025).

5. AI Adoption in the Public Sector. – URL: https://www.ictworks.org/wp-content/uploads/2025/03/ai-public-sector-survey.pdf (date of access: 29.09.2025).

6. Research Snapshot The future of Generative AI in policy work. – URL: https://www.unsw.edu.au/content/dam/pdfs/unsw-canberra/research/research-excellence/2025-06-the-future-of-generative-ai-in-policy-work.pdf (date of access: 29.09.2025).

7. Generative AI in Digital Government. – URL: https://dga.gov.sa/sites/default/files/2025-06/Generative%20AI%20in%C2%A0Digital%20Government-V1.0.pdf (date of access: 29.09.2025).

8. Почти 80 % компаний в Беларуси постоянно используют ИИ в маркетинге. – URL: https://belretail.by/article/pochti-kompaniy-v-belarusi-postoyanno-ispolzuyut-ii-v-marketinge (дата обращения: 29.09.2025).

УДК 004.056

Роль когнитивных ядер при решении прикладных задач

М. В. Стержанов
Белорусский государственный университет
информатики и радиоэлектроники, Минск
E-mail: sterjanov@bsuir.by

Введение

Понятие когнитивного ядра (КЯ) можно рассматривать как инженерную абстракцию минимальной функциональной конфигурации, которая принимает входные сигналы и целенаправленно преобразует их в действия, опираясь на восприятие, формализацию знаний, рассуждение и обучение [1]. В отличие от монолитных систем КЯ четко задает интерфейсы для подключения модулей памяти, поиска знаний, инструментов и исполнительных агентов, а также каналы прослеживаемости решений [1–4]. Такая модульность упрощает эволюцию систем: от узкоспециализированных решений к более универсальным платформам.

Практическая ценность КЯ проявляется в стандартизации подходов к диагностике, прогнозированию, планированию и управлению, что ускоряет разработку и повышает воспроизводимость прикладных решений [1, 3, 4].

Понятие когнитивного ядра и его свойства

Под КЯ понимается согласованная композиция алгоритмов и представлений, которая поддерживает замкнутый цикл «восприятие → интерпретация → принятие решения → действие → обратная связь → обучение» [1]. Ключевые свойства КЯ включают:

– многоуровневость представлений: от сенсорных признаков и латентных факторов к символическим и причинным структурам, что позволяет наращивать абстракции без потери связности [3, 5, 6, 7];

– адаптивность и перенос: способность оперативно перестраивать поведение при смене среды и целей, в том числе за счет обучения с подкреплением [4, 8, 9];

– обучаемость с разными типами сигналов: с учителем, без учителя, с подкреплением, по инструкциям и через инструментальные действия, что расширяет спектр применимости [4, 5, 8, 10];

– доступ к внешним ресурсам: подключение баз знаний, поисковых индексов, графов знаний и внешней памяти для обогащения контекста [3, 11–13];

– прослеживаемость и верифицируемость: контроль источников, оценка неопределенности и объяснимость, повышающие доверие к результатам [4, 7, 14, 15].

История развития подходов к когнитивным ядрам

1950–1970. Кибернетика и ранний символизм. Ранние системы опирались на четкие символические представления, продукционные правила и логический вывод, системно разделяя восприятие, память и рассуждение [1, 16].

1980–1990. Когнитивные архитектуры. Универсальные архитектуры (ACT-R, Soar) моделировали аспекты человеческого познания, четко различая долговременную и рабочую память, механизмы активации и внимания [17, 18].

1990–2015. Коннекционизм и глубокое обучение. Нейросетевые методы сформировали высокоэффективный перцептивный «фронт-энд», обеспечив масштабируемые представления и значительные прикладные результаты, но одновременно обострив вопросы объяснимости и контроля; существенный импульс придало глубокое обучение с подкреплением [5–9].

2015–н.в. Нейросимволические и гибридные системы. Современные КЯ объединяют глубинное обучение с явными структурами знаний и инструментами (поиск, символьные решатели, симуляторы) [2–4, 10, 11]. Ориентированные на поиск (retrieval) схемы связывают вероятностное предсказание с проверяемыми фактами [11], а агентные подходы интегрируют рассуждение и действие [10].

Принципы построения когнитивных ядер

Рассмотрим основные принципы, которые задают «конституцию» архитектуры [2]:

1. Нейрофункциональная аналогия. Каждый компонент системы проектируется, исходя не из точной биологической реализации, а из функциональной роли, эквивалентной соответствующему отделу головного мозга.

2. Гибрид нейросимволического ядра. Архитектура сочетает символическую память (граф знаний, который можно явно запрашивать и изменять) с нейронными компонентами (например, LLM или глубинные модели восприятия). Символический граф памяти служит каркасом знаний, а нейронные системы обеспечивают интуицию, обобщение и обработку «шумных» или неструктурированных данных. При решении задач обе

части взаимодействуют: память выдает релевантные факты, которые затем запускают нейронную процедуру рассуждения, и наоборот.

3. Радикальная модульность и специализация. Система разбивается на независимые агенты, каждый из которых отвечает за строго определенную функцию (память, планирование, эмоции, пространство, внимание и др.). Такая специализация позволяет улучшать и отлаживать каждый модуль независимо, избегая перегрузки одного узла всем функционалом. Модули обмениваются информацией по строго определенным интерфейсам и протоколам.

4. Иерархическое слоистое построение. Агентная система строится как слоистая архитектура: агенты нижнего уровня взаимодействуют с внешней динамикой (сенсорика, быстрые реакции), агенты среднего уровня интегрируют и агрегируют информацию, а верхний агент выполняет планирование, контроль и управление. Это позволяет абстрагироваться от «шума данных» и давать верхним уровням уже подготовленные, упорядоченные входы. Также обеспечивается механизм глобальных сигналов (например, сигнал тревоги), который может «пробиться» через слои к исполнительным агентам.

5. Всепроникающая наблюдаемость и способность к интроспекции. В архитектуре изначально предусмотрено журналирование всех сообщений между агентами, фиксирование внутренних состояний и метаданных. Кроме того, предусмотрены механизмы «самонаблюдения» – агент (или сам исполнительный агент) может время от времени анализировать журналы событий, выявлять аномалии, переосмысливать решения или корректировать поведение. Это делает систему объяснимой и позволяет вмешиваться в ее функционирование, а также служит механизмом регулирования и безопасности.

Данные принципы не существуют изолированно, а переплетаются – модульность сочетается с иерархией, наблюдаемость поддерживает гибридный дизайн, а нейрофункциональная аналогия направляет специализацию модулей.

Применение когнитивных ядер

Сводный анализ по когнитивным архитектурам фиксирует более 900 проектов в 84 архитектурах [4]. Рассмотрим некоторые важные области применения КЯ.

Робототехника и автономные системы

Автономные робототехнические системы предъявляют сразу несколько требований к когнитивной архитектуре:

– интеграция разнородных модулей: восприятие (камера, лидар, сенсоры), локализация и карты, планирование движения, исполнительные механизмы;

– многоуровневое управление: от нейронных политик для низкоуровневых реакций до символических планировщиков для долгосрочных задач;

– работа в реальном времени с вероятностной оценкой окружения и предсказанием поведения агентов;

– объяснимость и прослеживаемость решений (особенно в корпоративных или военных робототехнических системах);

– обучение с переносом через симуляторы и микромодели.

КЯ объединяет нейросетевые модули (CNN, трансформеры) для восприятия, алгоритмы SLAM и байесовское планирование для навигации, символический уровень для постановки целей и механизмы извлечения информации для подключения инструкций, регламентов и корпоративных баз знаний [3, 6–9, 11–13].

КЯ в робототехнике формирует иерархическую систему управления, где нижний уровень отвечает за восприятие и реакцию в реальном времени, средний уровень – за траектории и навигацию, а верхний – за постановку и объяснение задач.

В качестве примера рассмотрим работу автономного сервисного робота в складском комплексе. На первом шаге данные камер и лидаров подаются в сверточные и трансформерные модели, что дает роботу возможность распознать стеллажи, проходы, людей. Нейросетевой модуль классифицирует объекты: «человек», «коробка», «погрузчик». Затем алгоритм SLAM строит карту склада. Вероятностная локализация (например, Monte Carlo Localization) учитывает шум сенсоров. После этого система планирует движение к нужному стеллажу и корректирует траекторию в реальном времени при появлении препятствий. Высший уровень КЯ формулирует задачи: «найди стеллаж A23», «забери коробку», «доставь к пункту сортировки». Планировщик может использовать PDDL-подобное представление, где цели и предикаты описаны явно. Робот обращается к корпоративному хранилищу инструкций: «как безопасно перемещать коробку с опасным грузом». Внешняя долговременная память хранит историю предыдущих маршрутов и опыта (например, где чаще бывают загромождения) [11–13]. При выборе маршрута робот балансирует исследование (поиск новых путей) и эксплуатацию (использование уже известных безопасных маршрутов), применяя байесовское планирование [3]. Перед внедрением новых стратегий движения робот тестирует их в цифровом двойнике (симуляторе склада). Микромодель позволяет прогнозировать поведение людей и техники, чтобы снижать вероятность аварий [20].

КЯ сочетает нейросетевое восприятие с символическим планированием и поиском, что дает прозрачность и адаптивность. В отличие от традиционной схемы, которая жестко разделена на слои и плохо адаптируется к неопределенным условиям, КЯ динамически балансирует вероятностные и символические методы, позволяя роботу учиться и переиспользовать опыт. Известна «хрупкость» систем, основанных только на онтологиях, так как неожиданный объект ломает модель. КЯ интегрирует онтологию с нейронными перцептивными модулями и вероятностными фильтрами. КЯ нацелено на промышленное применение, включая объяснимость и соответствие корпоративным регламентам.

Медицина и клиническая поддержка решений

Клиническая практика требует от интеллектуальных систем:

– обоснованности и прозрачности (каждая рекомендация должна быть верифицирована источниками – руководствами, базами лекарственных взаимодействий, клиническими исследованиями);

– интеграции разнородных данных (визуальные данные МРТ/КТ, текстовые записи врачей, биохимические анализы, геномика, базы лекарств);

– калиброванных прогнозов риска (например, вероятность осложнений при выборе определенной терапии);

– причинно-следственного анализа, а не просто корреляций.

КЯ обеспечивает это за счет:

– машинного зрения (обработка снимков, выявление патологий);

– NLP-модулей (извлечение симптомов и заключений из текстов историй болезни);

– вероятностных графовых моделей (оценка риска осложнений);

– уровня извлечения информации (поиск в клинических руководствах, онтологиях, базах лекарственных взаимодействий [3, 4, 11]);

– решающего слоя с причинными представлениями и калиброванными прогнозами [7, 14, 15].

В качестве примера рассмотрим пациента с жалобами на боль в груди и одышку. На первом шаге машинное зрение анализирует ЭКГ и КТ-снимок грудной клетки с целью выявления признаков возможной ишемии и легочного застоя. NLP-модуль обрабатывает записи врача для извлечения симптомов «одышка при нагрузке», «боли в груди», «повышенное давление». Затем следует интеграция в граф знаний: Узел «ишемическая болезнь сердца (ИБС)» связывается с «боли в груди + ЭКГ-аномалии». Узел «сердечная недостаточность» связывается с «одышка + легочный застой». Модель риска рассчитывает вероятности: ИБС = 0,72, сердечная недостаточность = 0,55. Затем происходят поиск и обоснование: система запрашивает рекомендации ESC (European Society of Cardiology). Система сверяет

назначения с базами лекарственных взаимодействий (например, осторожность при одновременном применении бета-блокаторов и антиаритмиков). После этого система формирует сценарии. Назначить нагрузочный тест + эхокардиографию (с опорой на рекомендации ESC). Начать терапию низкими дозами бета-блокаторов (с учетом взаимодействий). Вероятность улучшения состояния в течение шести месяцев при соблюдении терапии – 0,65 (с интервалом доверия ±0,1). На заключительном шаге система сообщает врачу: «Рекомендуется нагрузочное тестирование, так как вероятность ИБС выше 70 %. Основание: ESC Guidelines, отчет по пациенту. Для назначения терапии учтено лекарственное взаимодействие.

В отличие от нейросетевых решений, которые могут диагностировать по снимку, но не объясняют решение, КЯ комбинирует зрение, текст, вероятностные модели и дает обоснование со ссылками. Экспертные системы (например, старые CDS) основаны на жестких правилах(«если-то») и не адаптируются к контексту. КЯ использует динамический граф знаний и вероятностные выводы, позволяя учитывать сложные комбинации факторов. Оно строит причинно-следственные модели и сценарные прогнозы, что повышает клиническую ценность рекомендаций.

КЯ связывает машинное зрение, обработку естественного языка и вероятностные модели риска; слой извлечения информации обеспечивает доступ к клиническим руководствам, онтологиям и базам лекарственных взаимодействий [3, 4, 11]. Решающий слой опирается на причинно-следственные представления и калиброванные прогнозы с верифицируемыми ссылками [7, 14, 15]. Исторически показательна интеграция поиска, онтологий и доказательств в проекте Watson [2].

Образование и персонализированное обучение

Когнитивное ядро в обучении обеспечивает:

– гибкую адаптацию: система динамически перестраивает траекторию обучения под конкретного учащегося;

– диагностику знаний: обновляет вероятности усвоения каждой темы на основе поведенческих и когнитивных индикаторов;

– объяснимость: педагог и ученик видят, почему именно предлагаются те или иные задания;

– интеграцию открытых источников: подключение учебных материалов, научных статей, видео и практических кейсов через модуль поиска;

– персонализированную мотивацию: эмоциональные и когнитивные сигналы используются для подбора стиля и уровня сложности заданий.

В качестве примера рассмотрим учащегося вуза, изучающего теорию вероятностей. На первом шаге проходит анализ ответов в тестах: система фиксирует не только правильность, но и время ответа, частоту ошибок

определенного типа. Система формирует графовую модель. Узел «условная вероятность» получает низкий уровень усвоения (p = 0,35). Узел «правило умножения событий» – высокий уровень (p = 0,82). Система обновляет вероятности после каждого ответа. Распознавание неверных стратегий: например, ученик постоянно путает условную вероятность с полной вероятностью.

Затем система формирует индивидуальный план обучения для данного учащегося. Например, вставить дополнительные задачи на условные вероятности, предложить задачу из реального мира («как оценить вероятность выигрыша в двухэтапной лотерее»).

Система обращается к внешним ресурсам: находит адаптированные видео с объяснением теоремы Байеса, подбирает практические примеры из открытых библиотек. Система демонстрирует повышенную объяснимость. Ученику говорится: «Мы даем тебе больше задач на условную вероятность, так как три из последних пяти ответов были ошибочными». Учителю сообщается: «Вероятность усвоения темы «условная вероятность» = 0,35, рекомендуем повторение».

Таким образом, в отличие от традиционных систем управления обучением (LMS) система отслеживает динамику когнитивного состояния, формирует вероятностную модель знаний и дает адаптивные рекомендации. В отличие от адаптивных систем обучения (например, ALEKS, Coursera), которые используют ограниченные статистические алгоритмы адаптации, КЯ интегрирует перцептивные модули, граф знаний и внешние ресурсы, позволяя выйти за рамки заранее подготовленного образовательного контента. ИИ-преподаватели на базе LLM могут отвечать на вопросы, но не хранят долговременную модель знаний ученика. КЯ поддерживает персонализированную долговременную память, отслеживает эволюцию знаний и адаптирует задания на горизонте недель и месяцев [1, 5, 11, 14, 15].

Финансовая аналитика и принятие решений

Финансовая сфера предъявляет жесткие требования к системам поддержки решений:

– прослеживаемость (каждое решение должно быть объяснено с указанием источников данных и логики вывода);

– калиброванные прогнозы (вероятности и риски должны быть количественно обоснованы, а не являться «черным ящиком»);

– мультидоменная интеграция (рыночные котировки, отчетность компаний, макроэкономические показатели, регуляторные публикации и даже новостные потоки).

Когнитивное ядро обеспечивает это через гибрид нейросимволического подхода:

– нейронные модули анализируют большие объемы неструктурированных данных (новости, отчетность на естественном языке, тональность комментариев);

– символическая часть (граф знаний) формирует причинно-следственные связи: «рост процентных ставок → снижение привлекательности долгосрочных облигаций», «ухудшение отчетности компании → рост вероятности снижения рейтинга»;

– временные модели и вероятностные графы отслеживают динамику факторов, связывая текущие изменения с вероятностью будущих сценариев.

Рассмотрим задачу оценки перспектив акций крупного банка в условиях изменения ключевой ставки. На первом шаге осуществляется сбор данных (нейронные агенты): интерес представляют потоки новостей об изменении денежно-кредитной политики, финансовая отчетность банка, сообщения регулятора. После этого система переходит к шагу интеграции и интерпретации на основе графовой памяти. Например, узел «ключевая ставка ↑» связан с узлом «стоимость фондирования ↑», узел «стоимость фондирования ↑» связан с узлом «снижение маржи банков». Третьим шагом является построение вероятностной модели. Для примера приведем три сценария с указанием их вероятности. Сценарий А: ставка повышена на 1 % → вероятность падения прибыли банка на 15–20 % (вероятность 0,65). Сценарий В: ставка без изменений → прибыль стабильна, рост акций в пределах 2–3 % (вероятность 0,25). Сценарий С: ставка понижается → рост прибыли банка на 10–12 % (вероятность 0,10). Четвертым шагом является формирование прогноза. Приведем пример отчета, созданного КЯ: «С высокой вероятностью (>60 %) текущая денежно-кредитная политика приведет к сокращению рентабельности банковского сектора. Основные факторы риска: рост стоимости фондирования и увеличение доли просроченных кредитов. Рекомендуется рассмотреть хеджирование позиций». Отметим обоснованность и прослеживаемость всех доводов. В отчете система указывает конкретные источники: протокол ЦБ, квартальный отчет банка, последние кредитные рейтинги, рыночные индикаторы.

Таким образом, в отличие от чисто статистических моделей КЯ формирует причинно-следственные объяснения, связывает данные с источниками и указывает вероятность сценариев. КЯ сочетает правила и обучение на данных, адаптируясь к изменяющимся рыночным условиям и интегрируя новые знания. КЯ использует динамическое моделирование и вероятностные графы, что позволяет предлагать стратегии, а не только отчетность [11, 14, 15].

Ограничения, риски и требования к верификации

Рассмотрим основные ограничения, препятствующие широкому применению КЯ в реальном секторе:

– галлюцинации и вымышленные факты. При использовании подсистемы извлечения информации КЯ может синтезировать «убедительные, но ложные» связи, если источники противоречивы или неполны [14, 15];

– сдвиги домена (domain shift). Модель, обученная на одних данных (например, медицинских исследованиях до 2022 г.), может ошибаться при интерпретации новых практик или лекарств [4];

– невоспроизводимость результатов. Стохастические компоненты (нейросети, случайные выборки в планировании) делают работу КЯ непредсказуемой: одно и то же задание может дать разные ответы [14];

– скрытая предвзятость (bias). Данные могут содержать систематические перекосы (например, недопредставленность женщин в клинических исследованиях или компаний из развивающихся рынков в финансовых данных) [4, 15].

Приведем примеры возможных рисков использования:

медицинский: неверный диагноз или рекомендация терапии из-за искаженных данных;

финансовый: некорректный прогноз рынка может привести к убыткам;

социальный: предвзятые алгоритмы обучения могут усиливать неравенство;

операционный: автономный робот может принять опасное решение (например, неверно оценить траекторию в реальной среде).

Сформулируем минимальный набор инженерных практик для верификации работы КЯ:

1. Обязательный доступ к проверяемым источникам. Retrieval проводится с указанием ссылок и номеров версий. Вводится запрет на использование неподтвержденных данных в критических решениях.

2. Явное моделирование неопределенности. Байесовские и причинные модели должны возвращать не только прогноз, но и доверительный интервал. Например: «вероятность осложнений = 0,68, интервал доверия ±0,12».

3. Аудит данных и версий моделей. Каждое решение должно быть воспроизводимо: запись версии модели, набора данных и параметров. Аудит позволяет выявить, была ли ошибка вызвана обновлением модели.

4. Тесты на исторических данных (off-policy) и ретроспективное тестирование стратегий (backtesting). Проверка на исторических данных перед реальным применением (например, стратегия в биржевой торговле должна пройти тест на данных 2015–2023 гг.). A/B-тесты позволяют сравнивать несколько стратегий на ограниченных выборках.

5. Контроль безопасности и приватности. Выполняется за счет шифрования на всех уровнях, анонимизации и псевдонимизации персональных данных и соблюдения международных стандартов и регуляций GDPR/HIPAA. Важными являются ограничение доступа к персональным данным в механизмах поиска, а также постоянный мониторинг и защита от утечек.

6. Механизмы остановки и деэскалации. Встроенные механизмы аварийной остановки (kill switch) позволяют немедленно прекратить работу системы при обнаружении ошибки, аномального поведения или угрозы безопасности. Приоритет оператора в случае конфликта – это правило, по которому любое решение человека-оператора имеет высший приоритет и автоматически отменяет действия системы, если возникает противоречие.

7. Человек в цикле. В сценариях высокой ответственности (медицина, оборона, финансы) решения должны подтверждаться экспертами [4, 14, 15].

8. Многоуровневая независимая верификация. Использование внешних независимых систем для проверки рекомендаций. Например, медицинское КЯ проверяется сразу двумя системами: собственной моделью и сторонним справочником.

Заключение

Когнитивные ядра формируют инженерный каркас для построения надежных интеллектуальных систем. Историческое развитие – от символических архитектур к гибридным и retrieval-усиленным – постепенно снимает противоречие между масштабируемостью и объяснимостью [1–16, 2, 4, 6, 9, 10, 11]. Для промышленного внедрения критичны верифицируемость, безопасность и управляемость знаниями [4, 7–15]. Перспективы развития связаны с интеграцией микромоделей, причинного вывода и внешней памяти в агентных контурах «рассуждение-и-действие» [3, 4, 10, 12–20].

Список использованных источников

1. Goodfellow, I. Deep Learning / I. Goodfellow, Y. Bengio, A. Courville. – Cambridge, MA : MIT Press, 2016.

2. Graves, A. Neural Turing Machines / A. Graves, G. Wayne, I. Danihelka // arXiv:1410.5401. – 2014. – URL: https://arxiv.org/pdf/1410.5401 (date of access: 24.09.2025).

3. Doshi-Velez, F. Towards a rigorous science of interpretable machine learning / F. Doshi-Velez, B. Kim // arXiv:1702.08608. – 2017. – URL: https://arxiv.org/abs/1702.08608 (date of access: 24.09.2025).

4. Kotseruba, I. 40 years of cognitive architectures: core cognitive abilities and practical applications / I. Kotseruba, J. K. Tsotsos // Artificial Intelligence Review. – 2018. – Vol. 53, no. 1. – P. 17–94. – DOI: 10.1007/s10462-018-9646-y.

5. LeCun, Y. Deep learning / Y. LeCun, Y. Bengio, G. Hinton // Nature. – 2015. – Vol. 521. – P. 436–444.

6. Attention Is All You Need / A. Vaswani, N. Shazeer, N. Parmar [et al.] // Advances in Neural Information Processing Systems (NeurIPS), Long Beach, CA, USA, 4–9 Dec. 2017. – Long Beach, 2017. – P. 5998–6008.

7. Koller, D. Probabilistic Graphical Models: Principles and Techniques / D. Koller, N. Friedman. – Cambridge, MA : MIT Press, 2009.

8. Human-level control through deep reinforcement learning / V. Mnih, K. Kavukcuoglu, D. Silver [et al.] // Nature. – 2015. – Vol. 518. – P. 529–533.

9. Retrieval-Augmented Generation for Knowledge-Intensive NLP / P. Lewis, E. Perez, A. Piktus [et al.] // Advances in Neural Information Processing Systems (NeurIPS), Vancouver, BC, Canada, 6–12 Dec. 2020. – Vancouver, 2020. – P. 9459–9474.

10. On the opportunities and risks of foundation models / R. Bommasani, D. A. Hudson, E. Adeli [et al.] // arXiv:2108.07258. – 2021. – URL: https://arxiv.org/abs/2108.07258 (date of access: 24.09.2025).

11. Royse, C. The Cognitive Core: An Integrated Cognitive Architecture / C. Royse, B. Caudill // ResearchGate. – 2025. – URL: https://www.researchgate.net/publication/392774960_The_Cognitive_Core_An_Integrated_Cognitive_Architecture (date of access: 02.10.2025). – DOI: 10.13140/RG.2.2.24838.08003.

12. Hybrid computing using a neural network with dynamic external memory / A. Graves, G. Wayne, M. Reynolds [et al.] // Nature. – 2016. – Vol. 538. – P. 471–476.

13. Ha D., Schmidhuber J. World Models // arXiv:1803.10122. – 2018. – URL: https://arxiv.org/abs/1803.10122 (date of access: 24.09.2025).

14. A survey of methods for explaining black box models / R. Guidotti, A. Monreale, S. Ruggieri [et al.] // ACM Computing Surveys. – 2018. – Vol. 51, no. 5. – Article 93.

15. ReAct: Synergizing Reasoning and Acting in Language Models / S. Yao, Y. Yang, Y. Cui [et al.] // Intern. Conf. on Learning Representations (ICLR), Kigali, Rwanda, 1–5 May 2023. – Kigali, 2023.

16. Newell, A. Unified Theories of Cognition / A. Newell. – Cambridge, MA : Harvard University Press, 1990.

17. Anderson, J. R. How Can the Human Mind Occur in the Physical Universe? / J. R. Anderson. – Oxford : Oxford University Press, 2007.

18. Russell, S. Artificial Intelligence: A Modern Approach / S. Russell, P. Norvig. – 4th ed. – Boston : Pearson, 2020.

19. Sutton, R. S. Reinforcement Learning: An Introduction / R. S. Sutton, A. G. Barto. – 2nd ed. – Cambridge, MA : MIT Press, 2018.

20. Pearl, J. The Book of Why: The New Science of Cause and Effect / J. Pearl, D. Mackenzie. – New York : Basic Books, 2018.

УДК 543.449+004.942

Использование искусственного интеллекта при подготовке технических специалистов в вузе

М. В. Стержанов✉, **А. А. Григорьев**
Белорусский государственный университет информатики и радиоэлектроники, Минск
E-mail: sterjanov@bsuir.by

А. А. Шимбалев
Белорусский государственный педагогический университет им. М. Танка, Минск
E-mail: shymbel@nobility.by

Введение

Современная эпоха цифровизации и переход к «Индустрии 4.0» кардинально трансформируют роль искусственного интеллекта (ИИ) в различных сферах человеческой деятельности, включая систему высшего образования. Для технических вузов, таких как Белорусский государственный университет информатики и радиоэлектроники (БГУИР), внедрение ИИ перестает быть лишь инновационной инициативой и становится стратегической необходимостью, направленной на подготовку конкурентоспособных специалистов, способных эффективно работать в условиях быстро меняющихся технологических реалий [1, 2].

Особую актуальность приобретает интеграция ИИ в преподавание фундаментальных дисциплин – физики, астрономии. Современные исследования подчеркивают, что ИИ-инструменты способны значительно расширить дидактические возможности обучения: сделать его более адаптивным, интерактивным, персонализированным и мотивирующим [3]. В то же время реализация этих потенциалов сталкивается с рядом вызовов, требующих комплексного подхода на уровне образовательной политики, методики и инфраструктуры.

В данной статье проводится анализ ключевых вызовов и перспектив применения ИИ в техническом образовании с акцентом на опыт БГУИР, а также рассматриваются возможности использования ИИ-технологий в преподавании физики и астрономии. Особое внимание уделяется международному контексту, в частности практикам внедрения ИИ в ведущих университетах Китайской Народной Республики.

Основные современные вызовы:

1. Обновление учебных программ в соответствии с темпами технологического прогресса. Образовательные программы должны оперативно адаптироваться к новым технологическим реалиям, включая методы и инструменты ИИ [1].

2. Подготовка преподавательского состава и развитие методической базы. Даже при наличии современных ИИ-инструментов их эффективное применение требует от преподавателей соответствующей цифровой компетентности и методической готовности к интеграции новых технологий в учебный процесс [4].

3. Этические, социальные и нормативные аспекты. Внедрение ИИ в образование требует учета вопросов прозрачности алгоритмов, конфиденциальности персональных данных обучающихся, а также справедливости и непредвзятости ИИ-моделей.

4. Технические и инфраструктурные барьеры. Для масштабного внедрения ИИ необходима развитая ИТ-инфраструктура: вычислительные мощности, облачные платформы, цифровые лаборатории и техническая поддержка.

5. Разрыв между теорией и практикой. Многие академические курсы по ИИ остаются преимущественно теоретическими, не обеспечивая студентов достаточным объемом практических навыков применения ИИ в реальных инженерных и научных задачах [2].

Рассмотрим текущее состояние дел:

Республика Беларусь

В стране активно развиваются инициативы по интеграции ИИ в высшее образование:

– в ряде вузов, включая БГУИР, уже внедрены курсы по машинному обучению, нейронным сетям и интеллектуальным системам;

– университеты налаживают сотрудничество с промышленными партнерами для предоставления студентам реальных кейсов и проектов;

– появляются пилотные проекты по использованию адаптивных образовательных платформ и аналитике учебного процесса.

Однако применение ИИ в преподавании фундаментальных дисциплин, таких как физика и астрономия, остается ограниченным, что подчеркивает необходимость дальнейших исследований и практических разработок в этой области [3, 5].

Китайская Народная Республика

Китай демонстрирует передовой опыт интеграции ИИ в STEM-образование:

– Tsinghua University (Пекин). Платформа XuetangX использует ИИ для персонализации курсов по физике и инженерии. В курсах по электро-

магнетизму внедрены интерактивные симуляции с ИИ-обратной связью, позволяющие визуализировать распределение полей в реальном времени.

– Zhejiang University. Система Smart Classroom отслеживает вовлеченность студентов и адаптирует темп лекций. В курсе «Электромагнитная теория» нейросетевые модели генерируют задачи разного уровня сложности на основе индивидуальной успеваемости.

– Shanghai Jiao Tong University. Лабораторные работы по электромагнетизму проводятся в AR/VR-средах с ИИ-поддержкой, что позволяет «наблюдать» невидимые физические поля. Система автоматически оценивает отчеты и предлагает рекомендации по улучшению.

– Национальная инициатива «AI+Education». Поддерживаемая Министерством образования КНР, эта программа финансирует пилотные проекты по внедрению ИИ в преподавание STEM-дисциплин. В ее рамках разработаны унифицированные ИИ-платформы для визуализации векторных полей, используемые в десятках университетов.

Особенности применения ИИ при обучении физике и астрономии в вузе

Физика и астрономия являются примерами фундаментальных дисциплин, формирующих основу инженерного и научного образования. Однако традиционные методы преподавания нередко сталкиваются с трудностями: низкой мотивацией студентов, ограниченными возможностями лабораторных занятий, недостаточной индивидуализацией процесса обучения.

ИИ предлагает инновационные решения для этих проблем, включая адаптивные платформы, виртуальные лаборатории, автоматическую проверку заданий и интеллектуальные системы поддержки студентов. Использование ИИ способно повысить эффективность и доступность физического образования, а также лучше подготовить студентов к современным вызовам науки и техники.

Использование ИИ в обучении физике и астрономии можно классифицировать по трем уровням:

образовательный – персонализация обучения, автоматизация проверки, интеллектуальные ассистенты;

научно-исследовательский – работа с большими массивами данных, анализ астрономических наблюдений, моделирование физических процессов;

методический – диагностика ошибок, визуализация, интеграция AR/VR, геймификация.

Далее рассмотрим их более подробно.

Адаптивное обучение

ИИ позволяет строить индивидуальные образовательные траектории. Это особенно важно в физике и астрономии, так как студенты приходят с разным уровнем подготовки. Сильные студенты получают более сложные задачи (например, по квантовой механике, общей теории относительности или астрофизическим моделям). Начинающие студенты могут работать с упрощенными моделями и визуализациями (например, законы Ньютона, движение планет). Кроме того, ИИ способен внедрять систему оценки знаний с весовыми коэффициентами. В такой системе простые задачи могут иметь меньший вес, а сложные – больший. Это позволяет более точно отражать реальный уровень компетенций студента; стимулировать сильных студентов к решению задач повышенной сложности; дать начинающим учащимся возможность накапливать баллы за базовые упражнения, постепенно переходя к сложным.

Таким образом, ИИ обеспечивает не только гибкость в построении траекторий, но и объективность в оценке знаний, поддерживая как базовый, так и углубленный уровни изучения.

Визуализация и моделирование

Традиционные лаборатории ограничены ресурсами: оборудование дорогое, некоторые эксперименты небезопасны или физически невозможны.

В физике ИИ управляет виртуальными лабораториями, где студенты могут «проводить эксперименты» с электрическими цепями, лазерами, ускорителями частиц. В астрономии создаются траектории движения тел на основе уравнений небесной механики, симуляции эволюции звезд, развития звездных и планетных систем, движения галактик, эволюционных процессов в масштабах всей Вселенной.

Интеллектуальные обучающие системы и ассистенты

ИИ выступает как виртуальный репетитор. В физике ассистент может объяснять шаги решения задач по механике или электродинамике, выявлять ошибки и предлагать подсказки. В астрономии ИИ может вести диалог с учеником, отвечать на вопросы о планетах, звездах, астрономических инструментах, космогонии, космологии, сопровождая объяснения визуализацией (например, 3D-моделью Солнечной системы или Галактики). Это снижает нагрузку на преподавателя и делает процесс более интерактивным.

Автоматизация проверки знаний.

В физике и астрономии часто требуется проверять большое количество задач, лабораторных отчетов и тестов. ИИ способен:

– проверять решения задач, включая уравнения и графики;

– анализировать ошибки и давать комментарии (например, «неверно применен закон Кулона»);

– оценивать тексты и отчеты, выделяя ключевые элементы (введение, методика, выводы).

Это позволяет преподавателю сосредоточиться на методической и исследовательской работе.

Обработка данных и исследовательский компонент

Физика и астрономия – это науки о больших данных: эксперименты на Большом адронном коллайдере, наблюдения телескопов Hubble, James Webb и TESS генерируют петабайты информации. В обучении физике студенты могут работать с реальными экспериментальными данными, используя ИИ для их анализа (например, спектры, сигналы с датчиков). В обучении астрономии ИИ помогает анализировать огромные объемы информации, что позволяет ученым совершать революционные открытия, например, в области поиска экзопланет благодаря сотрудничеству NASA и Google AI, изучающих данные телескопа Kepler, исследования далеких галактик, еле отличимых от фона изображения, и Сверхновых.

ИИ не только делает студентов обучающимися, но и добавляет «исследовательскую» составляющую, потому что они могут подключиться ко многим астрономическим онлайн-проектам, выставляющим для анализа огромные массивы информации по исследованию Солнца, поиску астероидов, комет и др.

Визуализация и дополненная реальность

Физические и астрономические процессы часто трудно представить. ИИ генерирует интерактивные визуализации: кривые светимости звезд, траектории частиц, гравитационные волны, модели звездных систем. В AR/VR-средах студенты могут «путешествовать» в космосе, наблюдать взаимодействие планет, видеть поля и силы в реальном времени. В квантовой физике ИИ помогает визуализировать волновые функции и вероятностные распределения. Это развивает пространственное и абстрактное мышление.

Геймификация и мотивация

ИИ поддерживает игровые подходы: в физике – соревнования по решению задач, эксперименты с виртуальными частицами; в астрономии – симуляции миссий («стань оператором космического телескопа», «управляй марсоходом»). Игровая динамика делает обучение увлекательным, снижает «страх перед формулами».

Диагностика когнитивных трудностей

ИИ способен выявлять когнитивные барьеры, где студенты чаще всего ошибаются (например, путают инерцию и силу или неправильно интерпретируют диаграмму Герцшпрунга – Рассела в астрономии), какие концепции требуют дополнительных пояснений. На основе этого формируются методические рекомендации для преподавателей.

Заключение

Использование ИИ в подготовке специалистов в вузах Беларуси на примере БГУИР является стратегическим направлением развития образования. Применение ИИ в физике, астрономии позволяет повысить эффективность обучения, адаптировать его под индивидуальные особенности студентов и приблизить учебный процесс к реальным задачам цифровой экономики. Для реализации потенциала необходимы повышение квалификации преподавателей, развитие инфраструктуры, создание методик и внедрение этических норм.

Список использованных источников

1. Рассел, С. Искусственный интеллект: современный подход / С. Рассел, П. Норвиг. – М. : Вильямс, 2021. – 1152 с.

2. Jordan, M. Machine learning: Trends, perspectives, and prospects / M. Jordan, T. Mitchell // Science. – 2015. – Vol. 349, no. 6245. – P. 255–260.

3. Shabanova, N. A. The possibilities of artificial intelligence in teaching physics / N. A. Shabanova, E. A. Meshcheryakova // Вестник Самарского университета. История, педагогика, филология. – 2021. – Т. 27, № 3. – С. 78–85. – URL: https://cyberleninka.ru/article/n/the-possibilities-of-artificial-intelligence-in-teaching-physics (дата обращения: 02.10.2025).

4. Жарков, А. Д. Искусственный интеллект в образовании: перспективы и риски / А. Д. Жарков // Вестник БГУИР. – 2022. – № 3. – С. 45–52.

5. Григорьев, А. А. Цифровые технологии как основа развития образовательной среды для дисциплины физика / А. А. Григорьев // Мультимодальность и экосистемный подход как тренды современного образования : сб. материалов Всерос. науч.-практ. конф., Армавир, 6 нояб. 2024 г. / Межрегиональный центр инновационных технологий в образовании, Армавирский гос. педагог. ун-т ; редкол.: Е. А. Дьякова, Л. Н. Горобец, И. А. Крутова. – Киров : Изд-во МЦИТО, 2025. – С. 77–80.

УДК 004.8:616-07:[616.12-008.3:616-053.2/.7]

Применение искусственного интеллекта для диагностики нарушений ритма сердца у детей с позиции врача

Н. В. Строгая✉, **А. В. Петрученя, Г. М. Батян**
Белорусский государственный
медицинский университет, Минск
E-mail: nata931994@gmail.com

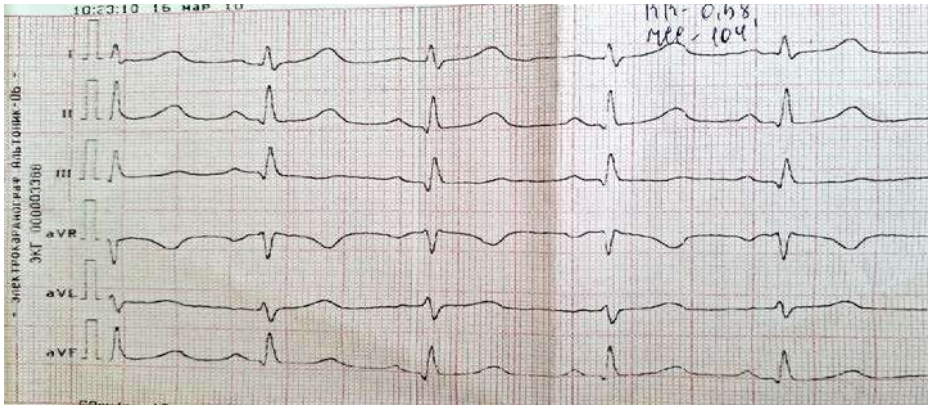
М. В. Соколова, В. С. Дроздова
УЗ «Городская детская инфекционная
клиническая больница», Минск, Беларусь

Введение

Искусственный интеллект (ИИ) находит применение в различных отраслях, включая здравоохранение. Анализ электрокардиограммы (ЭКГ) с использованием технологии машинного обучения (МО) представляется перспективным направлением, ориентированным на более персонализированном подходе к пациенту. ЭКГ – это одно из важных функциональных исследований, которое используется для диагностики различный заболеваний сердца у взрослых и детей и предоставляет большой объем данных, а также бесценную клиническую информацию для врачей различных специальностей.

Самым ценным станет внедрение ИИ в детскую кардиологию, что позволит упростить интерпретацию данных ЭКГ у рассматриваемой категории пациентов, повысив при этом точность диагностики патологических изменений у пациентов, находящихся в любом месте страны, до уровня заключений в учреждениях здравоохранения III-IV уровня.

ЭКГ с начала его использования в 1906 г. и до сегодняшнего дня остается самым распространенным методом функциональной диагностики в детской кардиологии. Автоматический анализ ЭКГ представляет собой непростую задачу, что связано с отсутствием четких границ между информативными фрагментами на реальных ЭКГ. Возрастные особенности электрофизиологии сердца у детей затрудняют автоматическое распознавание. Реальная ЭКГ содержит разного рода шумы, которые также оказывают влияние и вызывают искажение информативных фрагментов. Все это приводит к ошибкам в измерении диагностических признаков. На рисунке изображена реальная ЭКГ пациента в возрасте четырех лет.



Электрокардиограмма пациента четырех лет

При анализе ЭКГ врач учитывает и интерпретирует следующие значения зубцов, интервалов и сегментов: амплитуду и длительность зубца P, длительность интервала PQ, ширину комплекса QRS, длительность интервала QT, амплитуду зубца S, направления зубца T, которые имеют возрастные особенности, длительность интервала RR.

Стоит отметить, что в клинической практике используется расчет корригированного интервала QT для стандартизации с целью оценки риска развития синдрома внезапной сердечной смерти с использованием нескольких формул: Базетта (Bazett), Фреймингема/Саги (Framingham/ Sagie), Фредериции (Fridericia) и Ходжеса (Hodges).

При проведении анализа ЭКГ следует также учитывать особенности электрофизиологии сердца у детей, которые изменяются по мере взросления. Важно обеспечить высокое качество исходной записи ЭКГ за счет стандартной регистрации с 12 отведениями с правильным расположением электродов и временем записи, минимизации артефактов и шумов. Таким образом, для эффективного обучения моделей МО необходимо стандартизировать аннотацию по международным и национальным рекомендациям, привлекать экспертов для точной разметки и обеспечивать высокое качество исходных данных.

Важной технической проблемой остается удаление шумов на сигналах ЭКГ. Точность и надежность алгоритмов МО в значительной степени зависят от качества входных данных. Поэтому разработка более эффективных методов удаления шумов и предварительной обработки имеет решающее значение для повышения общей производительности аналитических систем ЭКГ. С другой стороны, материальные аспекты, такие как ограничения по времени и стоимости, являются серьезными препятствиями для широкого внедрения этой технологии. Например, вычислительная сложность некоторых моделей МО может мешать проведению анализа в реальном времени в условиях ограниченных ресурсов, из-за чего придется прилагать дополнительные усилия для оптимизации алгоритмов и обновления

аппаратного обеспечения. Кроме того, финансовые ограничения могут ограничивать доступ к современным инструментам анализа ЭКГ. Решение таких вопросов и работа над тем, чтобы сделать эти технологии доступными для всех, имеют большое значение для достижения равноправного доступа к медицинским услугам. К тому же повышенного внимания требуют вопросы конфиденциальности данных и приватности пользователей. Поскольку системы МО часто опираются на большие объемы персональных медицинских данных, существует обоснованная обеспокоенность по поводу возможного раскрытия и неправомерного использования конфиденциальной персональной информации. Поэтому надежное шифрование данных, безопасное хранение и прозрачная политика конфиденциальности необходимы для укрепления доверия пользователей и обеспечения выполнения нормативных требований по защите данных.

Разработанные ИИ-программы для ЭКГ у детей

На сегодняшний день в мире представлены десятки коммерческих и клинически верифицированных ИИ-программ для оценки изменений ЭКГ у детей. Каждая из этих программ может интегрироваться с аппаратным обеспечением или работать как облачный сервис для анализа ЭКГ в различных условиях. Некоторые публикации, представлены в таблице.

Публикации, основанные на использовании технологии ИИ в оценке ЭКГ у детей

Авторы	Год	Область применения в детской кардиологии	Применение технологии ИИ
Simona Aufiero [et al.] [1]	2022	Синдром удлиненного QT	Диагностика синдрома удлиненного интервала QT
S. Anjewierden [et al.] [2]	2023	ЭКГ-признаки дисфункции правого и левого желудочков	Выявление дисфункции правого и левого желудочков с использованием ЭКГ
J. Mayourian [et al.] [3]	2024	ЭКГ-признаки дисфункции левого желудочка и ремоделирования миокарда	Анализ ЭКГ для прогнозирования дисфункции левого желудочка и ремоделирования миокарда
J. Rahman, A. Brankovic, S. Khanna [4]	2024	Брадикардия у новорожденных	Использование ИИ для диагностики брадикардии у новорожденных
Y. S. Lee [et al.] [5]	2025	Прогнозирование врожденных пороков сердца с помощью ЭКГ	Использование ИИ для анализа ЭКГ с целью прогноза врожденных пороков сердца

Анализ ЭКГ с использованием методов машинного и глубокого обучения для диагностики удлиненного интервала QT (LQTS) был продемонстрирован в 2022 г. S. Aufiero [et al.] [1], где авторами были собраны ЭКГ,

обработаны, отфильтрованы и использованы для обучения моделей глубокого обучения.

Синдром удлиненного интервала QT (LQTS) представляет собой нарушение ритма сердца, характеризующееся удлинением интервала QT и аномальной морфологией зубца Т на ЭКГ. Для обучения модели, валидации и тестирования были взяты три отдельных подхода к обучению, в каждом из которых были использованы разные типы данных: первый метод – данные 12 отведений одной ЭКГ каждого пациента, второй – все полученные данные ЭКГ с 12 отведениями, третий – по одному отведению ЭКГ для обучения модели отдельно на каждом отведении. Этапы предварительной обработки данных включали фильтрацию, нормализацию и сегментацию сигналов ЭКГ. Авторы указывают на серьезные проблемы интерпретации детской ЭКГ, связанные с необходимостью учета возрастных изменений в физиологии сердца. Требуются возрастные референтные значения, поскольку нормальные пределы для младенцев заметно отличаются от нормальных значений для подростков. Модели ИИ должны быть обучены на возрастных данных и включать в себя механизмы учета изменений в развитии, чтобы избежать неправильной классификации и ошибочных диагнозов.

Использование ИИ включает в себя не только анализ ЭКГ с целью распознавания нарушений ритма сердца, но и анализ данных эхокардиографии (ЭХО-КГ), что позволяет дополнительно обратить внимание врача на структурно-функциональные изменения сердца [2]. Тем самым возможно использование ИИ для обнаружения скрытой патологии, которая у детей в силу возрастных особенностей может проявляться бессимптомно или неспецифически. В 2023 г. S. Anjewierden [et al.] [2] включили в исследование 10 142 пациента в возрасте до 18 лет, которым были сделаны трансторакальная ЭХО-КГ и ЭКГ покоя в 12 отведениях. На основании полученных данных с использованием сверточной нейронной сети, реализованной на платформе Keras с помощью языка программирования Python, была разработана модель, которая показала отличные результаты с AUC 0,93 (95 % ДИ 0,89–0,98) и точностью 89,0 % в диагностике дисфункции левого желудочка. При определении дисфункции правого желудочка AUC модели составила 0,90 (95 % ДИ 0,87–0,94) и продемонстрировала точность 81,0 % [2].

В отличие от исследования S. Anjewierden [et al.], модель J. Mayourian [et al.] [3] использовала не только данные 12 отведений ЭКГ, но и показатели фракции выброса левого желудочка, признаки гипертрофии, дилятации миокарда и оценку данных параметров по Z-score. При внешней валидации алгоритма значения AUC составили 0,94 для определения дисфунк-

ции левого желудочка, 0,84 и 0,87 – для гипертрофии и дилятации левого желудочка соответственно.

За последние несколько лет было проведено несколько исследований с использованием ИИ для диагностики нарушений ритма сердца у новорожденных. Следует отметить, что запись и интерпретация ЭКГ у новорожденных представляют определенные сложности. Это связано с возрастными особенностями, двигательными артефактами и невозможностью записи 12-канальной ЭКГ. В работе J. Rahman, A. Brankovic, S. Khanna [4] разработали новый подход для выявления брадикардии у новорожденных с использованием метода коррекции, чтобы снизить ложные тревоги (двигательные артефакты, судороги), который продемонстрировал AUC 0,98 и чувствительность 99,2 %. Использование коррекции позволило снизить количество ложных сигналов до 36,0 % по сравнению с отсутствием коррекции.

Одним из направлений использования ИИ в детской кардиологии является диагностика врожденных пороков сердца (ВПС) по данным ЭКГ. Хотя аускультация – это один из часто используемых методов диагностики ВПС, следует все же отметить, что имеется достаточно высокая эффективность в выявлении ВПС, связанных с поражением левых отделов сердца, и гораздо меньшая эффективность при выявлении ВПС, связанных с поражением правых отделов сердца, таких как дефекты межпредсердной перегородки, которые часто не вызывают шумов в сердце. ЭКГ позволяет определить некоторые типы ВПС, при которых изменения на ЭКГ проявляются раньше, чем патологические шумы в сердце.

В настоящее время существует настоятельная потребность в недорогостоящих и доступных методах скрининга для выявления некритических, но клинически значимых ВПС в периоде новорожденности и раннем детском возрасте. Модель, разработанная Y. S. Lee [et al.] [5], обучалась на данных ЭКГ 1035 пациентов в возрасте от 0 до 5 лет с пороками сердца с обогащением малого круга кровообращения, обеднением малого круга и обструкцией кровотока. При диагностике клинически значимых пороков с обогащением малого круга модель продемонстрировала AUC 0,85 (95 % ДИ 0,84–0,87) с чувствительностью 78,0 % и специфичностью 84,0 %, в детекции пороков с обеднением малого круга AUC составила 0,82 (95 % ДИ 0,77–0,86), чувствительность 69,0 %, специфичность 78,0 %.

Наличие различных нарушений ритма сердца диктует необходимость того, чтобы алгоритмы ИИ были способны выявлять разнообразные закономерности и различать условия с перекрывающимися характеристиками. Такие показатели, как чувствительность и специфичность, часто используемые для оценки эффективности модели с помощью AUC, могут переоценивать реальную полезность, когда прогнозируемое нарушение ритма серд-

ца встречается в популяции нечасто, но является клинически значимым. Следующая проблема связана с ограниченной доступностью и конфиденциальностью персональных данных.

Редкость определенных состояний и вариабельность регистрации ЭКГ на скорости записи 25 или 50 мм/с в разных учреждениях способствуют неоднородности данных. Небольшие размеры выборки могут привести к переобучению, когда модели хорошо работают на обучающих данных, но плохо на невидимых. Стратегии преодоления ограничений данных включают увеличение количества данных ЭКГ и совместный обмен данными между несколькими учреждениями. Нерепрезентативные наборы данных могут привести к алгоритмическим ошибкам, влияющим на диагностическую ценность. Внедрение инструментов ИИ в клиническую практику требует тщательного планирования и реализации. Необходимы эффективные программы для обучения врачей возможностям и ограничениям ИИ, повышения уверенности и правильного использования. Важно отметить, что ИИ должен служить вспомогательным инструментом для улучшения, а не замены клинических суждений, предлагая поддержку принятия решений и не диктуя осторожность.

Используя большой объем данных, алгоритмы ИИ распознают сложные закономерности и аномалии в сигналах ЭКГ, что способствует на современном этапе повысить точность диагностики. Кроме этого, анализ индивидуальных данных ЭКГ у детей позволяет предложить персонализированные рекомендации по лечению. Применение МО в анализе ЭКГ переворачивает представление о кардиологии, способствуя развитию объективного подхода, основанного на данных, и тем самым расширяя возможности как специалистов здравоохранения, так и пациентов для достижения лучших результатов.

Заключение

Применение ИИ для анализа ЭКГ у детей направлено на достижение трех основных целей: диагностики заболеваний сердца, определения точного типа заболевания сердца для конкретного пациента и прогнозирование вероятности развития осложнения выявленного заболевания и (или) других заболеваний сердца. Реализация технологии ИИ сможет улучшить методы анализа ЭКГ, заменив субъективные ручные подходы к интерпретации ЭКГ.

Список использованных источников:

1. A deep learning approach identifies new ECG features in congenital long QT syndrome / S. Aufiero, H. Bleijendaal, T. Robyns [et al.] // BMC Med. – 2022. – Vol. 20. – P. 162.

2. Detection of Right and Left Ventricular Dysfunction in Pediatric Patients Using Artificial Intelligence-Enabled ECGs / S. Anjewierden, D. O'Sullivan, K. E. Mangold [et al.] // J. Am Heart Assoc. – 2024. – Vol. 13. – e035201.

3. Pediatric ECG-Based Deep Learning to Predict Left Ventricular Dysfunction and Remodeling / J. Mayourian, W. G. La Cava, A. Vaid [et al.] // Circulation. – 2024. – Vol. 149. – P. 917–931.

4. Rahman, J. Machine learning model with output correction: Towards reliable bradycardia detection in neonates / J. Rahman, A. Brankovic, S. Khanna // Comput Biol Med. – 2024. – Vol. 177. – P. 108658.

5. Prediction of significant congenital heart disease in infants and children using continuous wavelet transform and deep convolutional neural network with 12-lead electrocardiogram / Y. S. Lee, H.-T. Chung, J.-J. Lin [et al.] // BMC Pediatr. – 2025. – Vol. 24. – P. 324.

УДК 004.032.26; 004.386

Обзор принципов построения аналоговых ускорителей вычислений для искусственных нейронных сетей

В. А. Сычёв
Объединенный институт проблем информатики
Национальной академии наук Беларуси, Минск
E-mail: vsychyov@robotics.by

Энергетические и технологические ограничения цифровых ускорителей нейросетевых вычислений

Успехи в технологиях искусственных нейронных сетей (ИНС) в последние годы достигнуты во многом благодаря развитию микроэлектроники. В то же время существует ряд факторов, которые влияют на применение ИНС как инструмента решения прикладных задач. В частности, по данным Международного энергетического агентства, затраты электроэнергии на работу центров обработки данных (ЦОД) во всем мире составили 415 ТВт/ч, т. е. около 1,5 % мирового потребления электроэнергии[1]. Для сравнения следует отметить, что выработка электроэнергии в Республике Беларусь в 2024 г. составила порядка 40 ТВт/ч[2]. При этом в рассматриваемой статистике не учтены портативные, носимые и встраиваемые устройства, которые также выполняют обработку данных, в том числе с использованием ИНС.

Следующим важным фактором называют размер элементов на интегральных схемах (ИС). Сейчас минимальное расстояние между затворами полевых транзисторов составляет порядка 50–57 нм для интегральных микросхем (ИМС), выполненных по технологии, обозначаемой «5 нм», где число служит маркетинговым обозначением поколения процесса, а не физического размера [1]. По мере приближения размеров элементов на кристалле ИМС к размеру атома кремния негативные эффекты миниатюризации проявляются во все большей степени [1, 2], а технологии производства усложняются. В настоящее время технологический процесс «5 нм» используется лишь двумя производителями микроэлектронной продукции.

[1]Энергосистемы как критический фактор развития ИИ-инфраструктуры. – URL: https://belnet.bsu.by/article/6252 (дата обращения: 26.08.2025).

[2]Производство электрической энергии. – URL: https://www.energo.by/contentobedineniya/proizvodstvo-elektricheskoy-energii (дата обращения: 26.08.2025).

При этом сбои в работе заводов и в логистических цепочках, наподобие тех, что имели место во время пандемии COVID-19 [3], могут оказать негативное влияние на развитие технологий искусственного интеллекта во всем мире.

Однако, несмотря на высокие затраты энергии и технологическую сложность оборудования, обучение и вывод ИНС на стационарных серверах и кластерах осуществляются успешно, что подтверждается высоким качеством работы новейших моделей ИНС. В то же время сейчас наблюдается тенденция по переносу вычислительных устройств ближе к источникам данных (так называемые граничные вычисления, англ. *edge computing*). В этом случае вывод и особенно обучение ИНС при ограниченных ресурсах вызывают технические сложности.

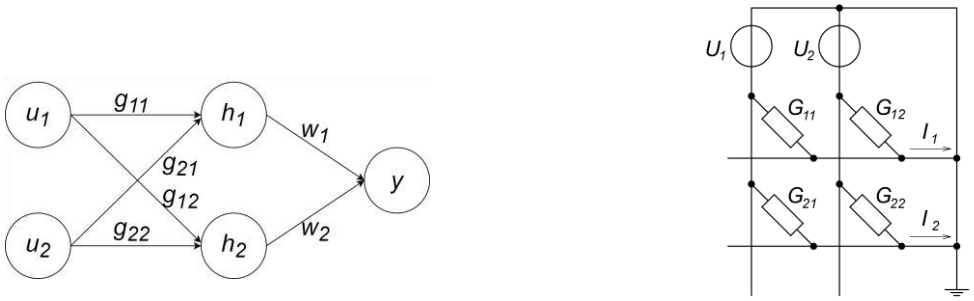
Одним из путей преодоления существующих ограничений как для стационарных систем, так и для портативных, носимых и встраиваемых устройств является разработка и применение приборов, специально спроектированных для выполнения узкого набора задач, называемых часто ускорителями. Предполагается, что специализированные устройства решают узкий набор задач многократно быстрее и с меньшими затратами энергии, чем универсальные компьютеры. Ранее ускорители в виде функционально законченных ИМС выполняли ресурсоемкие математические операции и назывались математическими сопроцессорами (спецпроцессорами). Позднее появились графические ускорители для задач, связанных с обработкой видео, и для вычислений, специфических для ИНС. В результате широкое использование ускорителей называется одной из новейших тенденций в развитии информационных и компьютерных технологий [4], что подтверждается в том числе архитектурой новейших ускорителей для генеративных ИНС.

Стремление получить энергоэффективные специализированные вычислительные устройства, которые могут быть произведены по доступным технологическим нормам, привело к появлению большого количества исследований путей ускорения вычислений, альтернативных существующим универсальным цифровым компьютерам. Одно из таких направлений обозначается в англоязычной литературе терминами unconventional, non-conventional [5], alternative, non-classical computing [6] и другими, аналогичными по смыслу, которые можно условно перевести как «альтернативные вычисления». Под альтернативными понимают квантовые, оптические, молекулярные, химические вычислительные средства [6] и в более общем смысле – *практику применения динамических систем (ДС) различной природы для выполнения вычислений с использованием их физических свойств* [5, с. 90, 7].

В качестве перспективной основы построения ускорителей вычислений рассматриваются также технологии аналоговой и гибридной, аналого-цифровой, вычислительной техники. С момента повсеместного внедрения цифровых машин аналоговые вычислительные устройства продолжили применяться для решения узкоспециальных задач, там, где именно аналоговые устройства могли продемонстрировать большее быстродействие и энергоэффективность, чем цифровые. Однако в последние годы аналоговые технологии нашли применение в области ускорения вычислений для ИНС. Далее будут рассмотрены принципы построения аналоговых ускорителей ИНС и наиболее значимые результаты, полученные исследователями в последнее время.

Принцип действия аналоговых ускорителей ИНС

С помощью аналоговых устройств реализуются схемы умножения и суммирования (в англоязычной литературе используется термин *multiply-accumulate, MAC*, или *multiply-add, MAD*), аппроксимации функций активации искусственных нейронов, а кроме того, элементы моделей биологических нейронов. Принцип действия аналогового ускорителя MAC показан на рисунке.



а) *б)*

Принцип действия аналогового ускорения работы ИНС:
а) структура ИНС с одним скрытым слоем; *б*) резистивная матрица

На рисунке, *а* представлена простейшая ИНС типа многослойного персептрона, действие которой описывается следующим образом:

$$\begin{aligned} h_1 &= f(g_{11}u_1 + g_{21}u_2 + b_1), \\ h_2 &= f(g_{12}u_1 + g_{22}u_2 + b_2), \\ y &= f(w_1h_1 + w_2h_2 + b_3), \end{aligned} \qquad (1)$$

где u_1, u_2 – входные сигналы ИНС, h_1, h_2 – выходные сигналы искусственных нейронов скрытого слоя, y – выходной сигнал ИНС, b_1, b_2, b_3 – смещения, $g_{11}, g_{21}, g_{12}, g_{22}$ – весовые коэффициенты нейронов скрытого слоя,

w_1, w_2 – весовые коэффициенты выходного нейрона, f – функция активации. Можно видеть из системы уравнений (1), что основные вычислительные затраты при работе ИНС связаны с выполнением многочисленных операций умножения и суммирования, которые могут выполняться аналоговыми вычислительными машинами (АВМ). Принцип работы аналогового устройства для выполнения этих операций проиллюстрирован на рисунке, *б*, где U_1 и U_2 – источники напряжений, величина которых определяется входными сигналами ИНС; $G_{11}, G_{21}, G_{12}, G_{22}$ – проводимости резистивных элементов, являющиеся аналогом весовых коэффициентов искусственных нейронов скрытого слоя $g_{11}, g_{21}, g_{12}, g_{22}$; I_1, I_2 – суммы токов, протекающих через первый и второй ряд резистивных элементов соответственно, которые определяются следующим образом:

$$\begin{aligned} I_1 &= U_1 G_{11} + U_2 G_{12}, \\ I_2 &= U_2 G_{21} + U_2 G_{22}. \end{aligned} \qquad (2)$$

Подстановка уравнений (2) в систему (1) дает

$$\begin{aligned} h_1 &= f(I_1 + b_1), \\ h_2 &= f(I_2 + b_2), \\ y &= f(w_1 h_1 + w_2 h_2 + b_3). \end{aligned} \qquad (3)$$

При практической реализации возможны различные варианты схемы на рисунке, *б.* В частности, последовательно с резисторами могут включаться транзисторы или диоды для предотвращения распространения тока по параллельным путям [9]. Существуют также различные способы представления отрицательных весовых коэффициентов, равно как и алгоритмы масштабирования весовых коэффициентов для исключения отрицательных значений [10].

Описываемый принцип был впервые предложен в работе [11] для построения ассоциативной памяти, однако из-за отсутствия технической возможности реализовать компактную управляемую резистивную матрицу практически значимые результаты были достигнуты много позже и в настоящее время стали основой технологий *вычислений в памяти*. Резистивная матрица, представленная на рисунке, *б*, может рассматриваться как запоминающее устройство, где проводимость G каждого резистора является аналогом некоей величины (в рассматриваемом примере – весового коэффициента). В англоязычной литературе для обозначения подобной памяти используется термин ReRAM (RRAM) – *resestive random-*

access memory, резистивная память со случайным доступом. Ее физическую основу составляют устройства, которые изменяют электрическую проводимость под воздействием управляющего сигнала и сохраняют изменения после снятия управляющего сигнала. Они называются мемристорами [12].

Мемристор представляет собой двухполюсник, демонстрирующий различные величины электрического сопротивления, не изменяющегося произвольно на значительном промежутке времени, который не потребляет энергию и может изменять свое сопротивление при приложении к полюсам коротких импульсов как тока, так и напряжения. Сопротивление подобного двухполюсника может быть измерено в любой момент времени с помощью испытательного тока или напряжения амплитудой, много меньшей, чем та, что использовалась при задании сопротивления. При этом испытательный сигнал подается на те же терминалы двухполюсника, что и сигнал управления. Отличительной чертой мемристора является петля гистерезиса на вольт-амперной характеристике.

Реализации аналоговых ускорителей вычислений для ИНС

Разработке ускорителя сетей глубокого обучения, позволяющего как обучать, так и использовать сеть для классификации, посвящена работа [13]. С помощью компьютерного моделирования авторами упомянутой работы были исследованы сети типа многослойного персептрона и сверточная ИНС, а также аналоговая реализация функций активации. В работе [14] на резистивной матрице размером 128×64, состоящей из пассивных элементов с изменяемой проводимостью, были продемонстрированы примеры спектрального анализа и обработки изображений. В статье [15] приводится обзор технологий, лежащих в основе резистивных устройств вычислений в памяти как для цифровых, так и для аналоговых устройств.

Помимо пассивных элементов с изменяемой проводимостью в качестве элементов памяти для вычислений могут применяться транзисторы с ловушкой заряда (англ. *charge-trap-transistors*, CCT), которые в настоящее время используются в ячейках энергонезависимых запоминающих устройств. В работе [16] представлена экспериментальная ИМС, содержащая массив CCT размером 784×784 элемента, производительностью 76,8 TOPS ($76{,}8\times10^{12}$ операций в секунду) при разрядности данных 8 бит и энергопотреблении 14,8 мВт на частоте 500 МГц.

В открытой печати есть сведения о применении аналоговых технологий в коммерческих разработках. Компания Aspinity[3] сообщает о разработке аналогового ускорителя ИНС, который потребляют менее 50 мка и поз-

[3]Aspinity. – URL: https://www.aspinity.com/ (date of access: 26.08.2025).

воляет ускорять работу сверточных, рекуррентных и других ИНС с числом параметров до 125 тыс. Основные задачи рассматриваемого устройства – использование ИНС для обработки данных от аналоговых датчиков, носимых на человеке, датчиков параметров окружающей среды. Существует возможность распознавания речи (отдельных слов) и изображений.

Ускоритель ИНС типа M1076 производства компании Mythic обладает производительностью 25 TOPS и позволяет хранить до 80 млн весовых коэффициентов для выполнения операций MAC без использования дополнительной памяти[4].

Ускоритель ИНС типа MX3 производства компании MemryX обладает вычислительной мощностью до 6 TOPS, позволяет использовать данные типа bfloat16 в функциях активации нейронов и применять 4, 8 или 16 бит для хранения весовых коэффициентов, причем для весовых коэффициентов разрядностью 8 бит ИМС позволяет хранить до 10,5 млн весов. Среднее энергопотребление одной ИМС составляет от 0,6 до 2,0 Вт. На основе описываемой ИМС строятся устройства форм-фактора M.2, которые включают в себя до 4 ИМС, в результате чего достигается производительность до 24 TOPS[5]. Следует отметить, что на сайте производителя отсутствуют прямые указания на использование аналоговых компонентов в рассматриваемых ИМС. В то же время коллектив разработчиков ИМС известен работами в области мемристоров и технологий вычислений в памяти [17].

Ограничения аналоговых ускорителей ИНС

Аналоговые вычислительные устройства имеют ряд технических недостатков. В частности, известно о влиянии на электронные схемы шума и неидеальности компонентов. В резистивных матрицах, построенных на мемристорах, проводимость ячеек может варьироваться из-за технологического разброса, дрейфа во времени и температурной нестабильности [14, 15]. Это приводит к ошибкам в весах нейронной сети, особенно критичным при обучении. В работе [14] отмечается, что при достигнутом авторами выходе исправных ячеек на уровне 99,8 % в массиве 128×64 точность записи проводимости ограничена: среднеквадратичное отклонение ошибки записи составило 6 мкСм при диапазоне проводимостей 100–900 мкСм. Подчеркивается, что флуктуации при чтении ограничивают эффективную разрядность системы ~7 битами, а в больших массивах точность снижается из-за паразитных сопротивлений проводников и токов утечки.

[4]M1076 Analog Matrix Processor. – URL: https://mythic.ai/products/m1076-analog-matrix-processor/ (date of access: 26.08.2025).

[5]MemryX. Products. – URL: https://memryx.com/products/ (date of access: 26.08.2025).

В работе [15] подробно анализируются физические источники неидеальностей, которые обусловлены стохастической природой формирования и разрушения проводящих филаментов в мемристорах. Кроме того, дрейф сопротивления вызван спонтанными атомными перестройками около филамента, что приводит к нестабильности весов после программирования. Указанные эффекты снижают точность вычислений и требуют применения корректирующих алгоритмов.

Не менее важно и то, что аналоговые архитектуры обладают ограниченным динамическим диапазоном [15, 16]. В отличие от цифровых систем аналоговые ускорители работают с непрерывными физическими величинами, которые затруднительно генерировать и измерять с высокой точностью. Это ограничивает эффективную разрядность представления весов и функций активации.

В работе [16] подчеркивается, что прямое представление 16-битных весов в одном мемристоре непрактично из-за требований к точности программирования и стабильности проводимости. Взамен используется квантование весов до 2 бит на ячейку и объединяются восемь ячеек для эмуляции 16-битного веса. Подобный подход требует специальной схемы сдвига и сложения и увеличивает аппаратные затраты, но позволяет обойти ограничения разрядности отдельной ячейки.

Динамический диапазон выходного тока в резистивной матрице ограничен как снизу (шумом и разрешающей способностью АЦП), так и сверху (максимальным током, который могут проводить соединения без значительного падения напряжения). Увеличение размера резистивной матрицы или точности весов экспоненциально повышает требования к АЦП.

Наконец, ограниченная разрядность затрудняет реализацию сложных нелинейных операций. Хотя наиболее распространенные функции активации могут быть аппроксимированы с помощью диодов или компараторов, их точность в аналоговом домене уступает цифровой реализации, особенно при малых значениях входного сигнала, где шум и нелинейности играют решающую роль.

Следующей важной задачей является реализация алгоритмов обучения ИНС непосредственно на аналоговой аппаратуре, что остается серьезной технической проблемой [5, 17]. Большинство аналоговых ускорителей ориентированы только на вывод, поскольку реализация алгоритмов обратного распространения ошибки средствами аналоговой электроники требует симметричных путей передачи сигнала «вперед» и «назад», что сложно или невозможно в пассивных резистивных сетях.

В обзоре [17] прямо отмечается, что, хотя массивы энергонезависимой памяти могут быть использованы как для прямого, так и обратного распространения ошибки за счет суммирования токов по строкам и столбцам, на

практике обучение сталкивается с ограничениями устройств. В частности, большинство элементов обладают асимметричной динамикой программирования: процесс увеличения проводимости часто является постепенным, тогда как процесс уменьшения проводимости – резким и плохо контролируемым, что нарушает симметрию градиентов, необходимую для корректной работы алгоритма обратного распространения ошибки, и приводит к падению точности при обучении. Также подчеркивается, что даже при успешной реализации обучения нелинейность и вариации устройств дополнительно ухудшают сходимость, особенно в сетях глубокого обучения. В результате большинство существующих аналоговых ускорителей либо используют предварительно обученные на цифровых платформах веса, либо ограничиваются локальными правилами обучения. Аналогичная проблема существует и в прочих системах, альтернативных архитектуре фон Неймана [5].

В-четвертых, масштабируемость и надежность аналоговых массивов ограничены паразитными эффектами, такими как утечки тока и кросс-токи [18, 19], а именно: при увеличении размера резистивной матрицы возникают паразитные эффекты, в первую очередь утечки тока по неактивным путям и кросс-токи между соседними линиями. Как отмечают в работе [18], даже при использовании селекторных транзисторов (1T1R-архитектура) паразитная емкость битовых линий и сопротивление соединений приводят к падению напряжения и искажению сигнала, особенно в крупных массивах. Это снижает точность операций умножения и суммирования и ограничивает практический размер массива резистивных элементов.

Кроме того, надежность мемристоров ограничена конечным числом циклов записи/стирания (endurance) [19]. В контексте аналоговых вычислений, где веса могут обновляться многократно, это становится критическим ограничением.

С точки зрения практической применимости аналоговых ускорителей следует отметить, что существует разрыв между лабораторными прототипами и промышленным производством. Как отмечают в статье [20], большинство демонстраций аналоговых ускорителей на основе мемристоров основаны на компьютерном моделировании, а не на реально изготовленных чипах. Кроме того, отмечается, что, несмотря на привлекательность аналоговых вычислений с точки зрения энергоэффективности, их внедрение в промышленность сдерживается нестабильностью параметров транзисторов и трудностями масштабирования аналоговых схем. Подчеркивается, что цифровые компьютеры сталкиваются с проблемами из-за захвата электронов в наноразмерных транзисторах, но аналоговые схемы еще более чувствительны к таким эффектам. Хотя нейроморфные системы, такие как Neurogrid, успешно используют смешанные аналогово-цифровые схемы в CMOS-

процессах (например, 180 нм), переход на передовые техпроцессы (например, 28 нм и ниже) требует полного перепроектирования аналоговых блоков из-за усиления стохастических эффектов и неоднородности устройств.

Выводы

Технология вычислений в памяти позволяет исключить этап обмена данными между устройствами обработки и хранения данных, характерный для архитектуры фон Неймана и приводящий к значительным потерям времени и энергии на передачу данных. Данное преимущество реализуемо как в цифровых, так и гибридных, аналого-цифровых, вычислительных устройствах. Однако в настоящее время подобная технология находится на начальном этапе развития. Таким образом, несмотря на очевидные преимущества аналоговых ускорителей в энергоэффективности и скорости выполнения операций умножения и суммирования, их практическое применение ограничено рядом взаимосвязанных проблем, начиная от физических неидеальностей устройств до отсутствия промышленно совместимых технологий массового производства. Основные особенности аналоговых ускорителей вычислений представлены в таблице.

Основные характеристики и ограничения аналоговых ускорителей вычислений для ИНС

Категория	Детали
Принцип действия	Выполнение операций умножения и суммирования через закон Ома и правила Кирхгофа в резистивных матрицах (кроссбарах); веса кодируются проводимостями; суммирование токов кодирует взвешенную сумму входов [11, 14]
Энергоэффективность	От <50 мкВт (Aspinity) до 14.8 мВт при 76,8 TOPS (ISAAC [16]); на порядки меньше цифровых решений при выводе
Производительность	Mythic M1076: 25 TOPS; MemryX MX3: 6 TOPS/микросхему (до 24 TOPS в M.2); ISAAC: 76,8 TOPS при 8 бит [16]
Емкость памяти весов	От 125 тыс. (Aspinity) до 80 млн (Mythic) и 10,5 млн при 8-битных весах (MemryX); веса хранятся в памяти
Технологическая основа	ReRAM/мемристоры [14]; транзисторы с ловушкой заряда (CCT) [16]; гибридные CMOS-аналоговые схемы (до 180 нм)
Точность и разрядность	Эффективная разрядность 6–7 бит из-за шума и дрейфа [14, 15]; 16-битные веса эмулируются через восемь 2-битных ячеек [16]
Обучение на кристалле	Крайне ограничено: асимметрия программирования нарушает градиентную симметрию; большинство решений – только вывод [17]
Ключевые ограничения	Дрейф проводимости, технологический разброс [14, 15]; паразитные токи и кросс-токи [18]; ограниченное число циклов записи/стирания [19]; сложность масштабирования [20]
Стадия внедрения	Преобладают моделирование и лабораторные образцы; коммерческие продукты (Mythic, Aspinity) нишевые, для edge-устройств [13, 14, 16]

Благодарность. Работа выполнена при поддержке гранта БРФФИ Ф23МЭ-030 «Эволюционные методы генерации оптимальных структур высокопроизводительных аппаратных ускорителей для реализации искусственных нейронных сетей на базе перепрограммируемых логических интегральных схем».

Список используемых источников

1. A Better Way to Measure Progress in Semiconductors. It's time to throw out the old Moore's Law metric. – URL: https://spectrum.ieee.org/abetter-way-to-measure-progress-in-semiconductors (date of access: 30.01.2025).

2. Kish, L. B. End of Moore's law: thermal (noise) death of integration in micro and nano electronics / L. B. Kish // Physics Letters A. – 2002. – Vol. 305, no. 3/4. – P. 144–149.

3. Frieske, B. The «semiconductor crisis» as a result of the COVID-19 pandemic and impacts on the automotive industry and its supply chains / B. Frieske, S. Stieler // World Electric Vehicle Journal. – 2022. – Vol. 13, no. 10. – P. 189.

4. IEEE International Roadmap for Devices and Systems. International Roadmap for Devices and Systems (IRDS™): 2023 Edition // IEEE. – 2023. – URL: https://irds.ieee.org/editions/2023 (date of access: 30.01.2025).

5. Nakajima, K. Reservoir Computing / K. Nakajima, I. Fischer. – Springer, 2021. – 463 p.

6. Teuscher, C. Unconventional computing catechism / C. Teuscher // Frontiers in Robotics and AI. – 2014. – Vol. 1. – P. 4.

7. Сычёв, В. А. Разработка метода классификации сигналов с использованием осциллятора Чжуа в рамках резервуарного подхода / В. А. Сычёв, А. М. Крот, Г. А. Прокопович // Журнал Белорусского государственного университета. Математика. Информатика. – 2023. – Vol. 1, no. 88. – P. 101.

8. A blueprint for precise and fault-tolerant analog neural networks / C. Demirkiran, L. Nair, D. Bunandar, A. Joshi // Nature Communications. – 2024. – Vol. 15, no. 1. – P. 5098.

9. In-memory computing for machine learning and deep learning / N. Lepri, A. Glukhov, L. Cattaneo [et al.] // IEEE Journal of the Electron Devices Society. – 2023. – Vol. 11. – P. 587–601.

10. Adaptive precision cnn accelerator using radix-x parallel connected memristor crossbars / J. Lee, J. K. Eshraghian, K. Cho, K. Eshraghian // arXiv preprint. – 2019. – URL: https://arxiv.org/pdf/1906.09395 (date of access: 24.09.2025).

11. Штейнбух, К. Автомат и человек: кибернетические факты и гипотезы / К. Штейнбух ; пер. с нем. – М. : Сов. радио, 1967.

12. Chua, L. Handbook of Memristor Networks / L. Chua, G. C. Sirakoulis, A. Adamatzky. – Springer Nature, 2019. – P. 1382.

13. Liu, X. Memristor crossbar architectures for implementing deep neural networks / X. Liu, Z. Zeng // Complex & Intelligent Systems. – 2022. – Vol. 8, no. 2. – P. 787–802.

14. Analogue signal and image processing with large memristor crossbars / C. Li, M. Hu, Y. Li [et al.] // Nature electronics. – 2018. – Vol. 1, no. 1. – P. 52–59.

15. Ielmini, D. In-memory computing with resistive switching devices / D. Ielmini, H.-S.P. Wong // Nature electronics. – 2018. – Vol. 1, no. 6. – P. 333–343.

16. ISAAC: A convolutional neural network accelerator with in-situ analog arithmetic / A. Shafiee, A. Nag, N. Muralimanohar [et al.] // Proc. of the 43rd Annual Intern. Symp. on Computer Architecture (ISCA '16). – Seoul, South Korea, 2016. – P. 14–26.

17. Burr, G. W. Neuromorphic computing using non-volatile memory / G. W. Burr, R. M. Shelby, A. Sebastian // Advances in Physics: X. – 2017. – Vol. 2, no. 1. – P. 89–124.

18. Gokmen, T. Acceleration of deep neural network training with resistive cross-point devices / T. Gokmen, Y. Vlasov // Frontiers in Neuroscience. – 2016. – Vol. 10.

19. Metal–oxide RRAM / H.-S.P. Wong, H. Lee, S. Yu [et al.] // Proceedings of the IEEE. – 2012. – Vol. 100, no. 6. – P. 1951–1970.

20. Efficient Processing of Deep Neural Networks: A Tutorial and Survey / V. Sze, Y.-H. Chen, T.-J. Yang [et al.] // Proceedings of the IEEE. – 2017. – Vol. 105, no. 12. – P. 2295–2329.

УДК 004.896

Гуманоидные роботы – воплощенный искусственный интеллект

Я. И. Туляков
ГУО «Гимназия № 13 г. Минска»

Введение

Робототехника и искусственный интеллект – это сегодня не только практические и научные изыскания, это новый взгляд на жизнь и прогнозирование образа будущей жизни людей: с чем будем жить, что использовать, как улучшится жизнь и какие возможности будут у людей. Объединение двух технологий – робототехники и искусственного интеллекта (воплощенный искусственный интеллект) – сформирует новый тип средств, так называемые человеко-машинные системы – интеллектуальные гуманоидные роботы, которые благодаря человекоподобным формам, подвижности и когнитивности смогут взаимодействовать и сосуществовать с людьми в мире, созданном для людей и самими людьми, что позволит улучшить жизнь и поднять цивилизационный уровень людей. Это переход от Индустрии 4.0 к Индустрии 5.0, которая отражает переход от акцента на экономическую ценность к акценту на общественную ценность, а также смещение акцента с благосостояния на благополучие, где стратегия, ориентированная на человека, «продвигает таланты, разнообразие и расширение прав и возможностей». Самый важный шаг, который это предполагает, – переход от восприятия людей как средств (например, как в человеческих ресурсах, которых заменят интеллектуальными роботами) к восприятию людей как целей, другими словами, в перспективе переход от людей, служащих организациям, к организациям, служащим людям. Гуманоидные роботы являются центральным звеном Индустрии 5.0.

Термин «гуманоидный робот» и цель «гуманоидной робототехники»

На сегодняшний день утвержденного определения термина «гуманоидный робот» не существует, однако под названием «гуманоидные роботы» (комбинация слов human и oid – человекоподобные роботы) понимаются роботы, которые выглядят как люди и имитируют человеческие движения и действия для выполнения различных задач. В идеале интеллектуальные гуманоидные роботы напоминают людей по форме (антропоморф-

ные черты, выражения), поведению (движения, жесты), восприятию и интеллекту (понимание, рассуждение, речь) и могут взаимодействовать с людьми естественным образом. Гуманоидная робототехника не является попыткой воссоздать людей. Цель состоит не в том, чтобы создавать машины, которые можно было бы ошибочно принять за настоящих людей, а в том, чтобы создать новый вид инструмента, принципиально отличающийся от любого, который использовался до сих пор, потому что он предназначен для работы с людьми и для людей. Гуманоиды будут взаимодействовать социально с людьми в типичных, повседневных условиях. Уже созданы и внедрены различные типы роботов (промышленные, сервисные, программные роботы) для выполнения утомительной, повторяющейся и иной работы, а гуманоиды – гуманоидные роботы – разрабатываются и создаются для безопасного сотрудничества и действия рядом с людьми или для замены людей, расширяя их возможности во всех сферах жизнедеятельности человека.

Гуманоидный робот, объединяющий классификации всех существующих роботов в одном роботе

До сих пор роботы разделялись следующим образом:

1. По воплощению – материальные роботы (роботы, функционирование которых требует физического воплощения), нематериальные роботы (роботы, которые не привязаны к конкретным физическим оболочкам).

2. По сферам применения – промышленные, сервисные (в том числе бытовые), боевые (в том числе роботы-солдаты), роботы для обеспечения безопасности, медицинские, исследовательские роботы, роботы для использования освоения космоса, поисково-спасательные роботы, пилотируемые, программные роботы (в том числе RPA-роботы, роботы (устройства) для средства отображения результатов и иные).

3. По конструкции – жесткие, мягкие, гибридные.

4. По типу – шарнирные, SCARA, цилиндрические, сферические, декартовы, параллельные (дельта), коллаборативные.

5. По функциональности – основные, вспомогательные, универсальные.

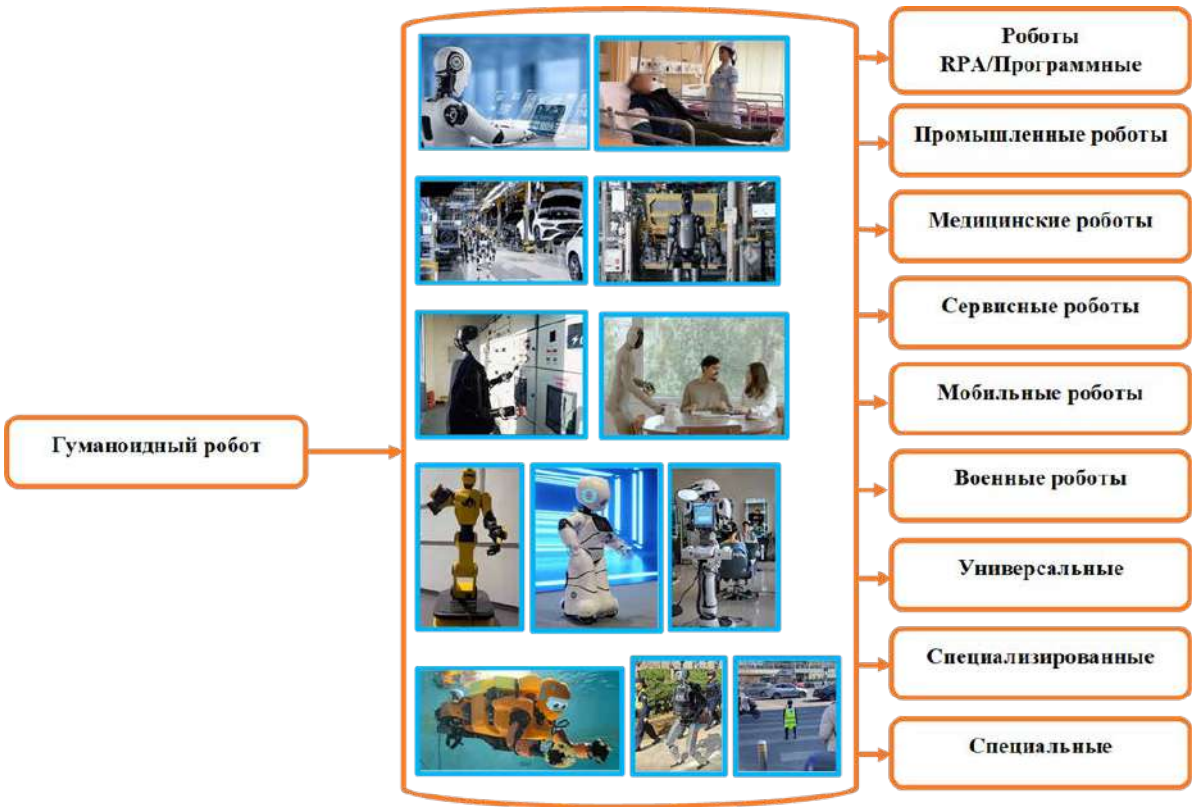
6. По управлению – переобучаемые (программируются жестко), обучаемые, самообучаемые.

Разделение роботов помогает сделать выбор, например, при их приобретении, дальнейшем применении и эксплуатации. Гуманоидные роботы в нашем понимании – это роботы, имитирующие действия человека и объединяющие все виды существующих роботов в одном роботе (рисунок), т. е. один гуманоидный робот может обладать всеми характеристиками и функциями, которыми по отдельности обладают известные нам роботы.

Итак, можно сделать ключевой вывод:

1. Гуманоидный робот может заменить любого из уже известных нам роботов независимо от применения и использоваться во всех сферах и секторах общественной жизни.

2. Гуманоидный робот – это универсальный робот, который занимает главенствующую (верхнюю) ступень технологического развития среди всех существующих роботов.



Гуманоидный робот

Возможности интеллектуальных гуманоидных роботов

Интеллектуальные гуманоидные роботы войдут в число «прорывных продуктов» после компьютеров, смартфонов, электромобилей. Они предназначены для имитации человеческого цикла «восприятие-познание-решение-исполнение». Будут ли гуманоидные роботы действительно работать? Будут ли они работать на производстве и обслуживать людей в различных сферах? Будут ли в будущем на нас работать гуманоидные роботы? Ответим на эти вопросы и сравним возможности гуманоидного (человекоподобного) робота с возможностями человека.

Мир создан для гуманоидного существа

Почему гуманоидные роботы смоделированы по образцу человеческого тела? Потому что людям комфортнее общаться и сотрудничать с себе подобными или на себя похожими. Потому что предприятия, города и целые страны построены для жизни гуманоидных существ и неважно – это человек или робот. Универсальность человеческого тела и гуманоидных роботов чрезвычайно ценна. Гуманоидные роботы, как и люди, могут преуспеть в неограниченном количестве задач и ситуаций. Они могут применять все инструменты, которые используют люди, манипулировать ими, осуществлять любые действия.

Новый тип роботов – гуманоидные роботы

В первую очередь гуманоидные роботы предназначены для работы, которую в противном случае выполняли бы люди: сначала на производстве и в логистике, затем в обслуживании и, наконец, в наших домах. В отличие от промышленных и складских роботов гуманоидные роботы могут работать в неструктурированной среде без необходимости вносить в нее изменения. Например, производственная линия – структурированная среда, а работа в обслуживании или по дому в значительной степени – неструктурированная. Работа гуманоидных роботов в структурированной среде и неструктурированной делает гуманоидов универсальными роботами, способными заменить человека в различных средах и сотрудничать с человеком при выполнении самых разных задач.

У гуманоидных роботов будут гораздо лучше развиты органы чувств, чем у человека

Люди могут выполнять все задачи с помощью шести чувств: глаза (зрение), уши (слух), язык (вкус), нос (обоняние), кожа (осязание, ощущение боли, температуры), вестибулярный аппарат (чувство равновесия и положения в пространстве, увеличение, ощущение веса). Однако люди способны воспринимать только небольшую часть реальности. У многих животных чувства гораздо более сильные. Орлы могут видеть в три раза острее, кошки могут видеть на 50 % лучше в темноте, собаки могут чувствовать запахи, по крайней мере, в 10 000 раз сильнее и слышать во много раз более высокие частоты, чем люди. Слоны могут чувствовать вибрации в земле на расстоянии нескольких километров. У некоторых животных есть дополнительные чувства, которых у людей нет вообще. Летучие мыши и дельфины используют эхолокацию. Змеи могут обнаружить кролика в темноте с помощью инфракрасного света. Пчелы, перелетные птицы и др. ориентируются с помощью так называемого магнитного чувства

(магнитное поле). Акулы могут воспринимать электрические поля живых существ, даже чувствовать их сердцебиение.

Все эти возможности можно передать гуманоидным роботам, например роботизированные глаза (компьютерное зрение), т. е. камеры с инфракрасным, ночным видением и специальные фильтры помогают видеть объекты ночью даже в тумане, что само по себе намного превосходит возможности человеческих глаз. Кроме того, роботы могут иметь круговое зрение на 360°, роботы с микрофонами слышат намного лучше, чем люди. Они могут иметь радары, лидары, ультразвуковые датчики, использовать так называемый «трибоэлектрический эффект» для искусственных пальцев своих рук, чтобы распознавать материал, к которому прикасаются, с точностью 90 %, что немыслимо для людей. Все это доказывает, что гуманоидные роботы могут обрабатывать значительно больше информации об окружающей среде, чем люди.

Гуманоидные роботы – эквивалент воплощенного искусственного интеллекта (ИИ)

Воплощенный ИИ относится к интеллектуальным системам, которые воспринимают и действуют на основе физического тела, взаимодействуя с окружающей средой для получения информации, принятия решений и выполнения действий, тем самым генерируя интеллектуальное поведение и адаптивность. Воплощенный ИИ не обязательно должен быть человекоподобным – гуманоидным. Гуманоидные роботы – это всего лишь одна из их физических форм. В центре внимания воплощенного ИИ находятся такие возможности, как восприятие, познание, обучение, принятие решений и взаимодействие с окружающей средой. Гуманоидные роботы благодаря воплощенному ИИ – это интеллектуальные роботы с человеческими внешними чертами и двигательными способностями. Они могут ходить на двух ногах, координировать руки и тело для функционирования и общаться с людьми. Итак, гуманоидные роботы получают «мозг» – ИИ. Другими словами, ИИ теперь получает тело.

Гуманоидные роботы многогранны и многофункциональны

В то время как люди обычно могут выполнять только одно или несколько действий, гуманоидный робот может выполнять все задачи, которым научился, или задачи, которые были в него заложены. Все, что нужно сделать, – это загрузить желаемый навык из многочисленных навыков в виде разработанного программного обеспечения (ПО) или применить технологию ИИ. Таким образом, робот сможет делать то, чему его научат. Если навык и действия не мешают друг другу, то роботы могут выполнять многозадачность, т. е. несколько сложных задач одновременно. Люди

только верят, что способны на многозадачность. Научно доказано, что это иллюзия.

Гуманоидные роботы всегда нацелены на высочайшее качество работы

В отличие от людей гуманоидный робот не отвлекается, не теряет концентрацию и мысли. У гуманоидного робота не бывает плохих дней. Гнев, грусть или иные эмоции не влияют на его работу. Качество работы гуманоидных роботов всегда будет высоким, точным и последовательным. Настолько высоким, насколько нельзя ожидать от человека.

Гуманоидные роботы ничего не забывают

Как много из изученного в школе, высшем учебном образовании и на различных курсах мы действительно помним? Как много из прочитанного в последней книге помним в деталях? Насколько уверены, что помним каждую деталь из описания того или иного процесса или инструкции? Таких вопросов можно задать очень много. В течение долгих лет люди должны учиться, чтобы получать знания и далее всегда и постоянно их наращивать. Гуманоидные роботы «рождаются» с приобретенными знаниями. Им не нужны годы учебы и, получив знания, они ничего не забывают, кроме устаревших знаний, которые должны будут «стереть» за ненадобностью. Поэтому знания и навыки гуманоидных роботов будут постоянно расти и обновляться. У людей нет никаких шансов угнаться за ними.

Гуманоидные роботы с ИИ учатся невообразимо быстро

Роботы-гуманоиды будут учиться у людей путем прямого наблюдения или по видео. Это называется имитационным обучением. Таким образом компания Tesla обучает ИИ для автономного вождения транспортных средств (у Tesla самый большой массив данных – практических видеоданных для обучения ИИ автономному вождению транспортных средств). Через текст в видео ИИ генерирует виртуальную симуляцию, что позволяет разработчикам моделировать даже самые сложные дорожные ситуации, т. е. создавать искусственные видео для обучения ИИ вождению транспортного средства. Это приближает разработчиков к поставленной цели – вождение ИИ транспортного средства должно быть в разы лучше и безопаснее, чем вождение транспортного средства человеком. Данная технология ИИ теперь используется в «мозгах» гуманоидных роботов и позволяет им работать на производственных предприятиях, в розничной торговле, в сфере общественного питания, в домашнем хозяйстве и пр.

Гуманоидные роботы будут обучаться сами по себе

ИИ и гуманоидные роботы обучались с помощью машинного обучения не один год, но теперь они все больше и больше учатся самостоятельно с помощью так называемого обучения с подкреплением (обучение с подкреплением (RL) – это метод машинного обучения, который обучает ПО принимать решения для достижения наиболее оптимальных результатов). Они самообучаются и им даже уже не нужны люди как учителя. Например, ИИ AlphaGo победил чемпиона мира по игре в го (го – это настольная игра для двух игроков, цель которой состоит в том, чтобы отгородить бóльшую территорию, чем у противника). Го намного сложнее шахмат. С 2019 г. у людей, играющих в го, не было ни единого шанса против ИИ. ИИ практически самостоятельно освоил этот навык без помощи людей. Человек становится гроссмейстером го за полжизни, за 30–40 лет, а AlphaGo стал супергроссмейстером всего за 40 дней без обучения со стороны человека.

Гуманоиды будут учиться друг у друга

Если один гуманоидный робот научится чему-нибудь, то у каждого другого гуманоидного робота в мире будут эти знания, потому что он загрузит в себя навык с платформы обмена навыками. Такая передача навыков от одного гуманоидного робота к другому происходит без многолетнего обучения, мгновенно. Можно предположить, что производители гуманоидных роботов ограничат то, чему может научиться гуманоид, и будут этот процесс контролировать. Возможности гуманоидных роботов с ИИ накапливаются. Невозможно вообразить ту скорость, с которой интеллектуальные гуманоидные роботы обмениваются информацией. Они способны передавать друг другу в глобальном масштабе огромные массивы своих навыков и знаний, которые приобретают ежеминутно, ежедневно и т. д., и все это при минимальных затратах по сравнению со стоимостью человеческого труда. У людей нет никаких шансов так быстро осваивать новые навыки.

Гуманоидные роботы будут настолько же разумны, насколько ИИ будет когнитивно разумен. Системы ИИ станут «мозгом» гуманоидных роботов и получат тело.

Гуманоидный робот физически сильнее, чем человек

Гуманоидный робот, имея металлический корпус и слаженную работу механизмов, будет в разы физически сильнее, чем человек.

Гуманоидные роботы могут быть использованы в опасных условиях

Роботам-гуманоидам не нужен воздух для дыхания, вирусы и бактерии не могут навредить им. Они могут продолжать работать в условиях

пандемии, в токсичном дыму, даже в высокорадиоактивных средах, если они построены соответствующим образом, и, конечно, роботы могут использоваться там, где риск потери человеческой жизни очень высок.

Гуманоидные роботы возьмутся за «недостойную работу»

Человеческие достоинства должны быть достаточной причиной, чтобы человека не заставляли делать скучную, грязную и опасную работу в современном мире только для того, чтобы заработать на жизнь, но можно ожидать, что роботы будут делать все это и в нашем понимании возьмутся за «недостойную» работу.

Гуманоидные роботы ломаются реже

Гуманоиды не болеют ни физически, ни психически. У них будет меньше несчастных случаев на работе, потому что они могут лучше и быстрее воспринимать окружающую обстановку. Если у них будут поломки, их заменят резервным роботом и отремонтируют.

Гуманоидные роботы могут работать гораздо дольше

Роботы не устают. Им, очевидно, не нужны перерывы и отпуска. Они не требуют пятидневной рабочей недели. Гуманоидам нужно прерывать свою работу только для зарядки. Это означает, что гуманоиды могут работать не менее 22 ч в сутки, семь дней в неделю и 360 дней в году. Это в четыре раза больше времени работы/часов, чем у людей (7920 ч вместо 1840).

Гуманоидные роботы могут работать быстрее

С развитием технологии гуманоидов гуманоидные роботы будут работать все быстрее. Человек не сможет конкурировать по скорости движения или выполнения действий с роботом. Итак, гуманоиды могут работать дольше и быстрее. Оба фактора говорят, что робот более чем в 10 раз имеет большую производительность, чем человек. Прежде всего это означает, что затраты на производство и услуги, а следовательно, и стоимость производимой продукции и услуг значительно упадут в среднесрочной перспективе. Это производительность, которая, в свою очередь, ведет к процветанию и качеству жизни.

Гуманоидные роботы не будут в дефиците

Во всем мире наблюдается дефицит кадров, особенно рабочих специальностей, и ситуация ухудшается. Все меньше и меньше людей готовы браться за скучную, грязную и опасную работу. Особенно в сфере производства, логистики, торговли и т. д. предприятия испытывают большие трудности с поиском достаточного количества рабочих. Гуманоиды будут браться за любую работу, и им неважно, она скучная, грязная или опасная. В будущем будет еще сложнее найти рабочих. Можно предположить, что

с серийным выпуском гуманоидов их стоимость не будет дорогой и она будет снижаться.

Гуманоидный робот-солдат

Гуманоидные роботы являются идеальными солдатами, и они могут быть солдатами в любой сфере.

Результаты

Ответим на три вопроса:

1. Будут ли гуманоидные роботы с их возможностями выполнять большую и растущую долю работы, которую раньше выполняли люди? Ответ «да».

2. Будут ли физические возможности гуманоидов достаточными для выполнения практически любой физической работы? Ответ «да».

3. Будут ли они достаточно разумны, чтобы делать все за нас? Здесь тоже ответ – однозначное «да».

Гуманоидные роботы будут достаточно когнитивно разумны, чтобы работать на людей и для людей. Гуманоидные роботы смогут достичь колоссальных физических и когнитивных показателей, что приведет к серьезным и революционным изменениям. Они найдут свой путь в нашей рабочей и жилой среде и кардинально изменят ее.

Прогнозное развитие интеллектуальных гуманоидных роботов у экспертов различно. Сценариев развития гуманоидных роботов предложено и опубликовано большое количество. Рассмотрим один из них.

Прогнозный сценарий развития гуманоидных роботов представлен крупнейшим финансовым институтом CITI – Global Perspectives & Solutions США. В таблице представлен общий количественный прогноз развития в мире рынка всех типов интеллектуальных автономных роботов – роботов, управляемых без человека, в период с 2024 по 2050 г. с шагом в один год.

Прогнозный сценарий развития всех типов автономных роботов, млн ед.

Названия	2024	2025	2030	2035	2050	CAGR, %
Автономный транспорт	27	34	126	380	1,858	17,4
Роботы-уборщики в домохозяйстве	286	326	541	793	1,188	5,3
Промышленные (территориальные) роботы-уборщики	1	2	6	14	25	10,6
Промышленные роботы	4,4	5,8	9	28	181	17,7
Гуманоидные роботы	0,01 (10 тыс. ед)	0,1 (100 тыс. ед)	1	13	1,1	50,1-60,7
БПЛА (дроны)	37	40	54	73	149	5,4

Окончание таблицы

Названия	2024	2025	2030	2035	2050	CAGR, %
Интерактивные роботы различной разновидности (роботы для общения)	0,5	1	6	18	71	20
Роботы-доставщики, или роверы	0,3	1	4	11	19	19,2
Роботы-демонстраторы (общепит, роботы-бармены, роботы, обслуживающие посетителей «с голоса» без участия человека-кассира)	0,1	0,7	3	10	15	17,2
Итого	**356,3**	**370,6**	**750**	**1,340**	**5,306**	

Прогнозируется, что к 2035 г. количество автономных интеллектуальных роботов в мире составит 1,340 млрд ед., из них гуманоидных роботов 12 млн ед., к 2050 г. – 5,306 млрд ед., из них гуманоидных роботов 1,1 млрд ед.

Заключение

Мир робототехники развивается беспрецедентными темпами, и гуманоидные роботы находятся на переднем плане технологического развития. Эти роботы – не просто машины, они являются аппаратным воплощением ИИ. Эксперты сходятся во мнении: 2024 г. стал началом эры «воплощенного искусственного интеллекта». Многие страны делают ставку на робототехнику – гуманоидных роботов – как на ключевой фактор технологического и экономического лидерства. Развитие технологий гуманоидных роботов является решением проблем глобального и беспрецедентного дефицита трудовых ресурсов, с которыми многие страны и секторы экономики столкнутся в предстоящие десятилетия. Согласно анализу численность трудоспособного населения будет сокращаться на 10–25 % в период с настоящего момента по 2050 г. В то же время спрос на трудовые ресурсы, особенно рабочих специальностей, будет расти, и гуманоидные роботы заполнят этот спрос, тем самым революционизируют различные сферы: от здравоохранения и образования до производства и развлечений.

Список использованных источников

1. Рассел, С. Искусственный интеллект. Современный подход / С. Рассел, П. Норвиг ; пер. А. В. Слепцов. – М. : Вильямс, 2021. – 704 с.

2. Wang, Y. Humanoid Robots 2025-2035: Technologies, Markets and Opportunities / Y. Wang // IDTechEx. – URL: https://www.idtechex.com/en/research-report/humanoid-robots/1093 (date of access: 24.09.2025).

3. Искусственный интеллект в современном мире // SAS Institute Inc. – 2025. – URL: https://www.sas.com/en_us/insights/analytics/what-is-artificial-intelligence.html (дата обращения: 12.10.2025).

4. Humanoid Robots Ready for the Masses: Report. – URL: https://www.iotworldtoday.com/robotics/humanoid-robots-near-mass-production-report (date of access: 24.09.2025).

5. Варанаси, Л. Основатель Figure AI Бретт Эдкок говорит, что скоро человекоподобных роботов будет столько же, сколько людей / Л. Варанаси // Insider Inc. – 2025. – URL: https://www.businessinsider.com/figure-ai-founder-brett-adcock-humanoid-robots-2025-7 (дата обращения: 12.10.2025).

УДК 004.032.26

Разработка нейросети для мобильного робота на базе LEGO Mindstorms EV3

Д. В. Черненко✉, А. Р. Ермолаева,
А. Ф. Сукорская, К. П. Швецов
Витебский государственный
технологический университет, Беларусь
E-mail: Chernenko203509@gmail.com

EV3 (Evolution 3) – это программируемый микрокомпьютер, являющийся «мозгом» роботов в системе LEGO Mindstorms Education EV3. Выпущенный в 2013 г. он имеет невысокие технические характеристики (процессор ARM9 300 МГц, 16 МБ флеш-памяти и 64 МБ оперативной памяти), однако позволяет использовать языки высокого уровня, такие как Java, C и Python. Несмотря на эти ограничения, его можно эффективно применять в исследовательских целях для решения задач прототипирования роботизированных систем, алгоритмов компьютерного зрения, навигации, искусственного интеллекта и машинного обучения.

Основная идея проекта разделена на несколько этапов и заключается в следующем:

1) обучение модели. На компьютере создается и обучается нейронная сеть, которая учится предсказывать правильные действия робота на основе данных с датчиков. Это способ был выбран как самый надежный, так как обучение нейросети требует больших вычислительных ресурсов;

2) развертывание модели. Обученная модель преобразуется в формат, понятный EV3, и загружается в память робота;

3) автономная работа. Робот использует модель для принятия решений в реальном времени без участия человека.

Для загрузки программ на языках высокого уровня модулем EV3 понадобится установить операционную систему, основанную на Debian Linux дистрибутиве – Ev3dev и работающую на совместимых с LEGO MINDSTORMS платформах, включая LEGO MINDSTORMS EV3. Скачать дистрибутив можно по ссылке: https://github.com/ev3dev/ev3dev/releases/download/ev3dev-stretch-2020-04-10/ev3dev-stretch-ev3-generic-2020-04-10.zip (355 МБ).

Запись образа осуществляется на карту памяти microSD или microSDHC объемом 2 ГБ.

Нужно вставить SD-карту в EV3 при выключенном питании, нажать центральную кнопку и удерживать ее до появления начала загрузки ev3dev.

Следующим шагом является выполнение настройки SSH, компилятора и C-библиотек. Для этого подключимся к ev3dev с помощью SSH. Подключение к SSH позволит выполнять команды на EV3 по сети, чтобы можно было развертывать код, изменять параметры и устанавливать инструменты.

Выполним установку компилятора C/C++ и проведем установку C-библиотек для работы с оборудованием EV3.

ev3dev-c – основная C-библиотека для управления моторами и датчиками LEGO EV3 под Linux. Библиотека ev3dev-c предоставляет методы работы с сенсорами, моторами и всеми ключевыми устройствами EV3. Для каждого типа сенсора данные читаются через устройства /sys/class/lego-sensor или через API-библиотеки.

Главной задачей проекта является разработка нейросети для навигации робота LEGO EV3. Цель проекта – разработать, обучить и внедрить в робот LEGO EV3 нейронную сеть, которая в реальном времени анализирует данные с ультразвуковых датчиков расстояния [2] и принимает решение о направлении движения, чтобы избежать столкновений с препятствием. Конструкция робота представлена на рис. 1.



Рис. 1. Внешний вид робота с двумя ультразвуковыми датчиками

Разрабатываемая модель представляет собой компактную нейросеть, предназначенную для интерпретации сенсорных данных и принятия решений о направлении движения робота. Ее архитектура включает два логических слоя обработки, что позволяет отнести ее к классу односкрытоточных нейросетей [1].

Модель принимает на вход два числовых параметра – расстояния до препятствий, полученные с левого и правого сенсоров робота (left, right).

Центральным элементом модели является скрытый слой, состоящий из двух нейронов. Каждый нейрон получает взвешенную комбинацию входных сигналов, преобразованную с помощью гиперболического тангенса (tanh) – активационной функции, обладающей симметрией и способной моделировать как положительные, так и отрицательные зависимости. Весовые коэффициенты (input_weights) на этом этапе играют ключевую роль в формировании внутреннего представления о пространственной конфигурации препятствий.

На финальном этапе скрытые активации поступают в один выходной нейрон, также использующий tanh в качестве активации. Это позволяет получить непрерывный выходной сигнал в диапазоне от -1 до 1, интерпретируемый как направление движения.

В рамках разработки интеллектуальной управляющей системы для мобильного робота была сформирована обучающая выборка в формате CSV, содержащая эмпирические данные, отражающие реакцию робота на препятствия. Данный набор представляет собой основу для обучения компактной нейросети, способной интерпретировать сенсорные сигналы и принимать решения о направлении движения (рис. 2). В датасете представлены разнообразные сценарии, включая симметричные и асимметричные конфигурации препятствий [3], что обеспечивает достаточную выборку для обучения модели.

```
Обучение нейросети...
Эпоха 0,    Ошибка: 0.352611 | Тест назад: [0.118652, 0.041269]
Эпоха 300,  Ошибка: 0.023810 | Тест назад: [-0.591572, -0.577210]
Эпоха 600,  Ошибка: 0.023433 | Тест назад: [-0.609288, -0.596101]
Эпоха 900,  Ошибка: 0.023290 | Тест назад: [-0.614738, -0.600896]
Эпоха 1200, Ошибка: 0.023137 | Тест назад: [-0.617107, -0.602372]
Эпоха 1500, Ошибка: 0.022985 | Тест назад: [-0.618686, -0.603316]
Эпоха 1800, Ошибка: 0.022836 | Тест назад: [-0.620056, -0.604260]
Эпоха 2100, Ошибка: 0.022688 | Тест назад: [-0.621392, -0.605283]
Эпоха 2400, Ошибка: 0.022540 | Тест назад: [-0.622757, -0.606379]
Эпоха 2700, Ошибка: 0.022389 | Тест назад: [-0.624164, -0.607524]
Достигнуто максимальное количество эпох. Финальная ошибка: 0.022233
```

Рис. 2. Отчет об обучении нейронной сети

Для отладки работы нейронной сети использовался режим симуляции робота. После выполнения режима симуляции в консоли программы выводится отчет, а затем и результаты выполнения тестовых ситуаций с пояснениями (рис. 3).

Реализация нейронной сети (2 входа, 2 скрытых нейрона, 1 выход) для управления роботом LEGO EV3 представлена на примере на языке Си (рис. 4).

```
== СИМУЛЯЦИЯ РОБОТА ===

Ситуация: Нет препятствий
Вход: [1.00, 1.00]
Выход: Левый мотор=0.72, Правый мотор=0.77
Действие: Движение вперед

Ситуация: Препятствие слева
Вход: [0.10, 1.00]
Выход: Левый мотор=-0.20, Правый мотор=0.90
Действие: Поворот налево

Ситуация: Препятствие справа
Вход: [1.00, 0.10]
Выход: Левый мотор=0.80, Правый мотор=-0.21
Действие: Поворот направо

Ситуация: Препятствие спереди
Вход: [0.20, 0.20]
Выход: Левый мотор=-0.39, Правый мотор=-0.41
Действие: Движение назад *** ЗАДНИЙ ХОД ***

Ситуация: Очень близко слева
Вход: [0.05, 0.80]
Выход: Левый мотор=-0.37, Правый мотор=0.77
Действие: Поворот налево

Ситуация: Очень близко справа
Вход: [0.80, 0.05]
Выход: Левый мотор=0.67, Правый мотор=-0.32
Действие: Поворот направо

Ситуация: Препятствия с обеих сторон
Вход: [0.10, 0.10]
Выход: Левый мотор=-0.63, Правый мотор=-0.61
Действие: Движение назад *** ЗАДНИЙ ХОД ***

Ситуация: Частичное препятствие спереди
Вход: [0.40, 0.40]
Выход: Левый мотор=0.23, Правый мотор=0.12
Действие: Неопределенное действие

=== ТЕСТ ДВИЖЕНИЯ НАЗАД ===
Вход: [0.10, 0.10] - препятствия с обеих сторон
Выход: Левый мотор=-0.63, Правый мотор=-0.61
Действие: Движение назад
```

Рис. 3. Отчет о выполнении симуляции тестовых ситуаций

```
float nn_predict(NeuralNetwork* nn, float* inputs) {
float hidden[2];  // Активации скрытого слоя
// Скрытый слой
for (int i = 0; i < 2; i++) {
hidden[i] = nn->hidden_biases[i];  // Начинаем со смещения
for (int j = 0; j < 2; j++) {
// Взвешенная сумма входов
hidden[i] += inputs[j] * nn->input_weights[j][i];}
hidden[i] = activate(hidden[i]);  // Применяем нелинейность}
// Выходной слой
float output = nn->output_bias;  // Смещение выхода
for (int i = 0; i < 2; i++) {
output += hidden[i] * nn->output_weights[i];  // Взвешенная сумма}
return activate(output);  // Финальная активация}
```

Рис. 4. Пример вычисления выхода нейросети на языке Си

Как видно из примера на рис. 4, в память робота загружается облегченная версия обученной нейросети в виде отдельного файла с обученными весами, который можно загрузить на робот. Эта версия идеально подходит для роботов с ограниченными ресурсами.

В результате проведенного эксперимента с моделью робота, созданного на базе LEGO Mindstorms Education EV3, можно сделать выводы, что даже слабому процессору ARM9, работающему с частотой 300 МГц и памятью 64 МБ, под силу работа с нейросетью, хотя и предварительно обученной. Таким образом, LEGO EV3 – это идеальный баланс между простотой использования и техническими возможностями. Он открывает путь в робототехнику для начинающих, при этом достаточно мощный для реализации серьезных проектов, включая работы с искусственным интеллектом, как в случае с нейросетями.

Список использованных источников

1. Хайкин, С. Нейронные сети. Полный курс / С. Хайкин. – М. : Вильямс, 2006. – 1104 с.

2. Николенко, С. И. Глубокое обучение. Погружение в мир нейронных сетей / С. И. Николенко, А. А. Кадурин, Е. О. Архангельская. – СПб. : Питер, 2018. – 480 с.

3. Джонс, М. Т. Программирование искусственного интеллекта в приложениях / М. Т. Джонс ; пер. с англ. А. И. Осипова. – М. : ДМК Пресс, 2018. – 311 с.

УДК 374.1

Обучение учащихся технологиям искусственного интеллекта как формирование компетенций будущего

А. В. Янковская
ГУО «Лидский районный центр творчества детей и молодежи», Беларусь

Введение

В современном мире искусственный интеллект (ИИ) становится ключевым инструментом, оказывающим значительное влияние на дополнительное образование детей и молодежи. В Республике Беларусь необходимость подготовки квалифицированных кадров в области ИИ становится все более актуальной, поскольку современный рынок труда требует специалистов, готовых работать с новыми технологиями и разрабатывать инновационные решения.

Обучение учащихся технологиям ИИ на уровне дополнительного образования играет главную роль в формировании у них необходимых компетенций и навыков. Это важно не только для обеспечения предполагаемой карьерной траектории будущих специалистов, но и стимулирования интереса к научной, исследовательской и проектной деятельности, что особенно востребовано в пятилетку качества [1]. Необходимы новые методические подходы и учебные материалы, которые помогут педагогам дополнительного образования внедрять современные технологии в учебный процесс, сделать его более интерактивным и увлекательным [2].

В Лидском районном центре творчества детей и молодежи кружок по программированию на Python для старшеклассников становится эффективной платформой для изучения основ ИИ и подготовки молодежи к дальнейшему обучению в этой области. Программирование на Python благодаря своей простоте и мощным библиотекам для работы с данными и машинным обучением выступает идеальной основой для знакомства учащихся с технологиями ИИ. Занятия в кружке предоставляют уникальную возможность развивать критическое мышление, творческий подход к решению задач и навыки работы с данными [3].

Цель данного исследования состоит в изучении методов и подходов обучения учащихся технологиям ИИ с целью формирования у них необходимых компетенций для успешной профессиональной деятельности в бу-

дущем, анализе существующих программ и разработке рекомендаций по внедрению технологий ИИ в сферу дополнительного образования.

Задачи исследования:

– проанализировать существующие образовательные программы в области ИИ для выявления лучших практик и методов обучения;

– определить ключевые компетенции, которые должны быть сформированы у учащихся в процессе обучения технологиям ИИ;

– разработать методические рекомендации по внедрению технологий ИИ в образовательный процесс для дополнительного образования;

– построить модель обучения с использованием интерактивных и проектных методов ИИ для более глубокого усвоения материала;

– провести эксперимент по применению предложенных методов обучения на практике и оценить их эффективность;

– исследовать влияние технологий ИИ на мотивацию учащихся к обучению и их профессиональную ориентацию.

Обзор источников по проблеме

Исследователи Ян Гудфеллоу, Йошуа Бенджио и Аарон Курвилль [4] описывают многочисленные эффективные методы и подходы к обучению ИИ, которые охватывают как теоретические, так и практические аспекты этой быстроразвивающейся области. Авторы акцентируют внимание на различных архитектурах нейронных сетей, таких как полносвязные сети, сверточные и рекуррентные нейронные сети, которые имеют свои уникальные достоинства в зависимости от типа решаемой задачи. Например, сверточные нейронные сети успешно справляются с задачами, связанными с изображениями, благодаря своей способности извлекать пространственные признаки, в то время как рекуррентные сети более эффективны при работе с последовательными данными, такими как текст. Освещается концепция обучения с учителем и без учителя, объясняя, как эти два подхода могут быть использованы для достижения успешных результатов в машинном обучении. Это понимание важно для выборки подхода, соответствующего конкретной задаче, а также для того, чтобы студенты могли осознанно применять эксперименты с размеченными и неразмеченными данными. Авторы придают особое значение оптимизации функции потерь и обсуждают различные методы, такие как градиентный спуск и его варианты, в том числе Adam и RMSprop. Данные методы помогают эффективно минимизировать потери модели, что принципиально для ее обучения. Кроме того, концепции регуляризации (L1 и L2 нормы или метод дропаута) помогают предотвратить переобучение моделей, что делает их более надежными и способными обобщать информацию на новых данных.

В книге содержится множество примеров и задач, что помогает учащимся закреплять теорию, улучшает их способность продуктивно применять методы глубокого обучения на практике.

Макс Тегмарк [5] исследует последствия развития ИИ для человечества и предлагаемые подходы к его обучению, а также подчеркивает, что ИИ является не просто инструментом, а новой формой жизни, которая требует переосмысления взаимодействия с технологиями. Он утверждает, что обучение технологий ИИ не должно ограничиваться только техническими аспектами, а важно включать и гуманитарные науки, что позволит будущим специалистам осознанно подходить к разработке алгоритмов и систем ИИ, которые будут учитывать потребности и ценности людей. Тегмарк предлагает интегрировать проекты и исследования в учебные планы, способствуя тому, чтобы студенты могли не только теоретически осваивать технологии, но и разрабатывать реальные приложения, решающие социальные или экологические проблемы. Это даст возможность учащимся видеть значимость своих навыков и действий. Автор также поднимает вопрос о важности социальной и этической ответственности, активного участия широкой общественности в обсуждении и принятии решений о том, как будет развиваться ИИ. Все указанное значимо для формирования нового поколения специалистов, способных создавать и использовать технологии, которые будут применяться в образовании.

Основная часть

Технологии ИИ широко используются в различных отраслях, начиная от финансов и медицины и заканчивая производством и транспортом. Важность определения ключевых компетенций, которые должны быть сформированы у учащихся в процессе обучения технологиям ИИ, заключается в том, что эти навыки и знания (программирование, аналитика данных, работа с большими объемами информации, машинное обучение) позволят создавать и адаптировать алгоритмы и модели, а также применять их на практике для решения различных задач. Помимо этого наличие указанных компетенций по ИИ способствует развитию критического и логического мышления учащихся, решению вопросов и инноваций не только в области технологий, но и в любой другой сфере деятельности, где требуется анализ данных, разработка стратегий и принятие решений на основе информации.

В процессе обучения технологиям ИИ на кружке Python у учащихся должны быть сформированы следующие ключевые компетенции:

1. Технические навыки программирования. Необходимо освоить основы языка Python, включая структуру данных, алгоритмы, а также специфические библиотеки для работы с ИИ (NumPy, Pandas, TensorFlow

и PyTorch), что является базой для дальнейшего изучения более сложных тем.

2. Понимание основ искусственного интеллекта. Важно знать основные концепции и методы ИИ, включая машинное и глубокое обучение, обработку естественного языка, анализ данных и др.

3. Аналитическое мышление и решение проблем. Требуется развивать способность анализировать данные, искать закономерности и разрабатывать алгоритмы для решения конкретных задач, что включает в себя умение ставить правильные вопросы и выбирать подходящие методы для их решения.

4. Критическое мышление. Принципиально развивать навыки критического анализа, чтобы осмысленно оценивать различные подходы и методы ИИ, включая их ограничения и социальные последствия.

5. Командная работа и сотрудничество. Важно развивать навыки работы в команде, включая коммуникацию, совместное принятие решений и умение делегировать задачи, поскольку многие проекты требуют взаимодействия с другими специалистами

6. Этическая грамотность. Существенно понимать этические аспекты применения ИИ, включая вопросы конфиденциальности, ответственности и предвзятости, что поможет разрабатывать технологии, которые учитывают интересы общества.

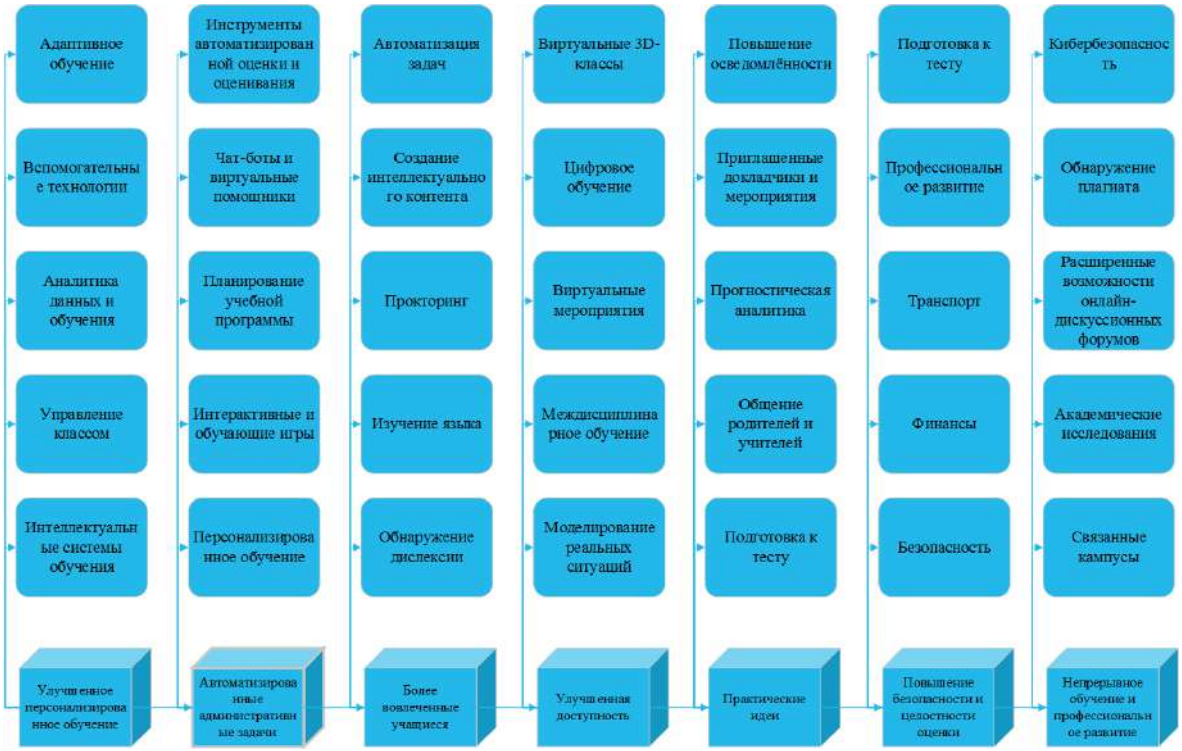
7. Адаптивность к изменениям. Развитие ИТ-технологий и ИИ стремительными темпами требует готовности к постоянному обучению и адаптации к новым инструментам, подходам и методам.

8. Практические навыки проектирования систем ИИ. Необходимо владеть навыками проектирования, создания и тестирования ИИ-систем, что включает в себя умение разрабатывать прототипы и проводить эксперименты с различными моделями.

9. Нарративные навыки и представление результатов. Важно уметь четко, доступно объяснять свои результаты и выводы заинтересованным сторонам.

Формирование указанных компетенций позволит учащимся не только успешно работать с технологиями ИИ, но и осознанно использовать их в будущем, учитывая как технологические, так и социальные аспекты.

ИИ производит революцию в образовании, предлагая передовые инструменты, которые улучшают преподавание и обучение – от персонализации опыта до оптимизации административных задач. Построим матрицу-модель (рисунок) обучения с использованием интерактивных и проектных методов ИИ для более глубокого усвоения материала.



Матрица-модель обучения с использованием интерактивных и проектных методов ИИ

ИИ как инструмент стремится трансформировать процессы обучения, предоставляя индивидуализированные, интерактивные и адаптивные образовательные возможности, которые значительно влияют на формирование внутренней мотивации учащихся и их стремление к карьерному росту.

Исследования показывают [6], что внедрение ИИ в образовательную среду способствует созданию более интересного и увлекательного контекста обучения. Виртуальные обучающие среды, основанные на алгоритмах машинного обучения, позволяют полнее учитывать индивидуальные особенности и предпочтения учащихся, что, в свою очередь, повышает их вовлеченность и удовлетворенность процессом обучения. Поскольку студенты получают возможность исследовать темы на своем уровне сложности и в индивидуальном темпе, такая адаптивность создает условия для их активной познавательной деятельности. В итоге учащиеся становятся более заинтересованными в обучении, так как видят прямую связь между усилиями и получаемыми результатами. Тем не менее повышение мотивации связано не только с персонализированным подходом, но и с внедрением ИИ в контекст профессиональной ориентации.

Технологии ИИ могут анализировать большие объемы данных о карьерных путях и тенденциях на рынке труда, предоставляя учащимся фактические данные о востребованных навыках и будущих профессиях. Эти данные позволяют студентам более осознанно принимать решения о своей

будущей карьере, сопоставляя личные интересы и способности с актуальными требованиями работодателей.

Дополнительно технологии ИИ внедряют новые форматы обучения, такие как смешанное (blended) или онлайн-обучение, которые дают возможность учащимся не только определять время и место для обучения, но и активно участвовать в процессе, а конструирование обучающих стратегий делает их более целеустремленными и сосредоточенными на достижении долгосрочных образовательных целей.

Влияние ИИ на мотивацию учащихся к обучению и их профессиональную ориентацию формируется в условиях социальной динамики и технологических изменений. Ожидание значительных преобразований на рынке труда, связанных с автоматизацией и внедрением новейших технологий, стимулирует учащихся развивать навыки, которые обеспечат их конкурентоспособность и изменят отношения к процессу обучения, а также определят будущую профессиональную идентичность.

Заключение

ИИ оказывает всеобъемлющее влияние на требования к профессиональным компетенциям, и обучение технологиям ИИ становится необходимым для подготовки квалифицированных кадров. Использование языка программирования Python в программах дополнительного образования детей и молодежи позволяет активно вовлекать их в процесс изучения технологий ИИ и разрабатывать практические проекты, что значительно повышает интерес и вовлеченность учащихся. Обучение технологиям ИИ должно включать в себя не только технические навыки, но и междисциплинарные знания, включая математику, логику и методы научного исследования, что содействует формированию более широкого кругозора у учащихся. Чтобы успешно реализовывать образовательные процессы, педагогам требуется поддержка в виде методических материалов и инновационных подходов для внедрения программ обучения технологиям ИИ. При освоении новых технологий открываются новые карьерные возможности, что способствует развитию креативного потенциала и инициативы учащихся. Таким образом, организация обучения учащихся ИИ в рамках дополнительного образования представляет собой стратегически важную инвестицию в будущее молодежи. Это поможет подготовить новое поколение специалистов, способных не только адаптироваться к современным вызовам, но и активно участвовать в их решении, формируя общество, основанное на инновациях и прогрессе.

Список использованных источников

1. О пятилетке качества : Указ Президента Респ. Беларусь от 17 янв. 2025 г. № 31 // Национальный правовой Интернет-портал Республики Беларусь. – URL: https://pravo.by/document/?guid=12551&p0=P32500031 (дата обращения: 23.09.2025).

2. О приоритетных направлениях научной, научно-технической и инновационной деятельности на 2026–2030 годы : Указ Президента Респ. Беларусь от 1 апр. 2025 г. № 135 // Национальный правовой Интернет-портал Республики Беларусь. – URL: https://pravo.by/document/?guid=3871&p0=P32500135 (дата обращения: 23.09.2025).

3. Концепция развития инженерного образования в Республике Беларусь на период до 2035 года : постановление Совета Министров Респ. Беларусь от 15 мая 2025 г. № 264 // Национальный правовой Интернет-портал Республики Беларусь. – URL: https://pravo.by/document/?guid=12551&p0=C22500264 (дата обращения: 23.09.2025).

4. Гудфеллоу, Я. Глубокое обучение / Я. Гудфеллоу, И. Бенджио, А. Курвилль; пер. с анг. А. А. Слинкина. – 2-е изд., испр. – М. : ДМК Пресс, 2018. – 652 с.

5. Tegmark, M. Life 3.0. Being Human in the Age of Artificial Intelligence / M. Tegmark. – New York : Alfred A. Knopf, 2017. – 364 p.

6. 39 примеров использования искусственного интеллекта в образовании [сайт]. – URL: https://onlinedegrees.sandiego.edu/artificial-intelligence-education (дата обращения: 23.09.2025).

Содержание

Введение 3

Абрамова Е. Н.
К вопросу о разработке понятия искусственного интеллекта в целях правового регулирования 5

Адамовский Е. Р., Богуш Р. П.
Обнаружение дыма на видеопоследовательности с использованием искусственной нейронной сети архитектуры трансформер 13

Андренко К. В., Головко В. А., Касперович С. А., Крощенко А. А.
Нейросетевая модель профориентационной диагностики: разработка веб-приложения 22

Бабоян Э., Дявго Я. А., Качан А. С.
Стратегии формирования национальных экосистем ИИ: сравнительный анализ ведущих стран мира 30

Бегунков В. И., Ковалев М. Я.
Классификация займа с помощью методов машинного обучения 42

Беляцкий Н. П.
Квантовая загадка искусственного интеллекта 48

Бибило П. Н., Романов В. И.
Технологически независимая оптимизация функций многозначной логики 55

Бранцевич П. Ю.
Автоматизация принятия решений при анализе состояния сложных объектов и ситуаций 63

Венгеренко В. В., Лукашевич М. М.
Мультимодальный анализ изображений для диагностики аномалий яичников с использованием машинного обучения 74

Владыкина Т. В., Постарнакова Ю. А.
Интеллектуальные мультиагентные системы 82

Владымцев В. Д.
Методы защиты информационных систем, использующих Large Language Models 91

Владымцев В. Д., Бахмат В. К., Крячев Е. В.
LLM в информационных системах: систематизация рисков и контрмер 100

Воеводин В. В., Лукашевич Н. В., Говин А. А.
Специализированная языковая модель как фундамент для сохранения уникальной социокультурной идентичности 109

Войтов И. В., Штепа В. Н., Дубина А. В.
Подходы интеллектуальной ресурсной аналитики промышленно-коммунальных объектов 115

Генрихсен С. Н., Седун А. М.
Классификация интеллектуальных инструментальных средств для автоматизации бизнес-процессов предприятия 122

Гончар А. В., Фурс К. В., Тузиков А. В., Андрианов А. М.
Применение генеративных нейронных сетей для дизайна новых ингибиторов белка KasA M. Tuberculosis 130

Гордей А. Н., Панащик Р. С., Святощик М. И., Стрельченок О. А., Ткаченко В. В., Шумилин А. Г.
Тего-семантические основания искусственного моделирования интеллектуальной деятельности 138

Денисов А. А.
Управление экспериментальным нейроэлектрофизиологическим оборудованием с применением нейросетевых языковых моделей 148

Дисковец Ю. М.
Искусственный интеллект в образовании: новые возможности для педагогической деятельности учителя 153

Жалова Д. А., Дыдо О. В., Хохлов В. А., Бакунович А. А., Журавлева Е. А.
Автоматическая система распознавания голосовых сигналов птиц, основанная на технологиях искусственного интеллекта 158

Жданович Е. С., Макренкова В. А., Алексеева Е. А.
Суверенный искусственный интеллект как основа цифрового будущего Беларуси 172

Исаев М. П.
Архитектура адаптивной образовательной среды на основе гибридных баз знаний и генеративных моделей 181

Капустин В. В.
Разработка электронной базы данных «Автопедии», интерактивных методов подготовки оценщиков, поврежденных дорожных транспортных средств и возможности использования искусственного интеллекта 185

Ким Т. Ю.
Программный комплекс управления роботизированными системами на основе технологии цифровых двойников, методов генетической оптимизации и машинного обучения 192

Ковалёва Н. Л., Ковалёв А. П.
Трансформация цифровой грамотности в ИИ-компетентность как ключевой ресурс адаптации преподавателя в цифровой образовательной среде 200

Косарева А. А., Снежко Э. В., Ковалев В. А.
Обнаружение узловых образований на КТ-изображениях легких на основе методов глубокого обучения 207

Кохнович Р. О.
Статистический анализ применения искусственного интеллекта для обнаружения таблиц в документах PDF 216

Курто Я. А.
Принципы обработки персональных данных системами искусственного интеллекта: содержание и современное состояние правового регулирования 221

Лабоцкая А. А.
Роль искусственного интеллекта в юриспруденции 228

Лапуцкий А. И., Сидорцов А. В.
Цифровая трансформация государственных органов на базе искусственного интеллекта 233

Лойко А. И.
Этические аспекты использования агентов искусственного интеллекта 241

Ляховец О. Н.
Пути повышения качества образования с применением технологий искусственного интеллекта 246

Прохоров Д. И.
Обучение учителей математики использованию технологии искусственного интеллекта для разработки учебно-методических материалов 254

Пунчик В. Н., Дорожкин К. В.
Перформативные аспекты внедрения искусственного интеллекта в образовательное пространство Беларуси 264

Резниченко И. А., Акиншева И. В.
Анализ данных для принятия решений в конструкторской подготовке 273

Романов В. И., Бибило П. Н.
Экспериментальное сравнение алгоритмов схемной реализации функций многозначной логики 279

Слесарева М. М., Зеновко Е. С., Макогон А. А., Кухаревич Г. С., Хохлов В. А.
Технологии синтеза речи в аудиогидах: проблемы и перспективы 287

Смышляева Ю. А.
Практика разработки ИИ-ассистента для коммуникационной системы высшего образования как элемент реализации государственной стратегии цифровой трансформации Республики Беларусь 300

Степанов В. А.
Использование генеративного искусственного интеллекта в практике работы государственных органов и организаций 309

Стержанов М. В.
Роль когнитивных ядер при решении прикладных задач 317

Стержанов М. В., Григорьев А. А., Шимбалев А. А.
Использование искусственного интеллекта при подготовке технических специалистов в вузе 328

Строгая Н. В., Петрученя А. В., Батян Г. М., Соколова М. В., Дроздова В. С.
Применение искусственного интеллекта для диагностики нарушений ритма сердца у детей с позиции врача 334

Сычёв В. А.
Обзор принципов построения аналоговых ускорителей вычислений для искусственных нейронных сетей 341

Туляков Я. И.
Гуманоидные роботы – воплощенный искусственный интеллект 352

Черненко Д. В., Ермолаева А. Р., Сукорская А. Ф., Швецов К. П.
Разработка нейросети для мобильного робота на базе
LEGO Mindstorms EV3..... 363

Янковская А. В.
Обучение учащихся технологиям искусственного интеллекта
как формирование компетенций будущего 368

Научное издание

IV МЕЖДУНАРОДНЫЙ ФОРУМ IT-АКАДЕМГРАДА
«ИСКУССТВЕННЫЙ ИНТЕЛЛЕКТ В БЕЛАРУСИ»

Доклады

Ответственный за выпуск С. С. Мойсейчик

Подписано в печать 27.11.2025. Формат 60×84 1/16. Бумага офсетная. Ризография.
Уч.-изд. л. 26,2. Усл. печ. л. 21,5. Тираж 30 экз. Заказ 14.

Издатель и полиграфическое исполнение:
государственное научное учреждение «Объединенный институт проблем
информатики Национальной академии наук Беларуси».
Свидетельство о государственной регистрации издателя, изготовителя,
распространителя печатных изданий № 1/274 от 04.04.2014.
Ул. Сурганова, 6, 220012, Минск.