



РЕГИОНАЛЬНЫЙ ФОНД
научно-технического развития Санкт-Петербурга (РФ НТР СПб)

ИНН 7825670231, КПП 780201001, ОГРН 1037843019697
Юридический адрес, фактический адрес: 194156, Россия, Санкт-Петербург, пр. Энгельса, д. 27, корп. 5, литера А
Тел.: +7 (812) 294-2511 E-mail: fntr@fntr.ru Сайт: <https://fntr.ru>

СОГЛАСОВАНО

Проректор по научной работе
Университета ИТМО
доктор технических наук, профессор


В.О. Низифоров
31 октября 2023 г.


УТВЕРЖДАЮ

Исполнительный директор
Регионального фонда научно-технического
развития Санкт-Петербурга


И.В. Гладких
31 октября 2023 г.


АНАЛИТИЧЕСКИЙ ДОКЛАД

о состоянии, перспективах развития областей науки и технологий, а также рынков
продукции (товаров, услуг), развивающихся с использованием результатов научной
и (или) научно-технической деятельности в Российской Федерации и за рубежом

Санкт-Петербург, 2023 г.

СОДЕРЖАНИЕ

ВВЕДЕНИЕ	5
1 Анализ состояния и основных тенденций и перспектив развития сферы науки и технологий по направлению 20а Стратегии научно–технологического развития Российской Федерации	7
1.1 Искусственный интеллект в компьютерном зрении	7
1.1.1 Искусственный интеллект в компьютерном зрении для детекции и идентификации объектов.....	7
1.1.2 Детекция и идентификация объектов в беспилотном транспорте.....	7
1.1.3 Детекция и идентификация объектов в промышленности	9
1.1.4 Детекция и идентификация объектов в здравоохранении.....	10
1.2 Искусственный интеллект в компьютерном зрении в робототехнике	11
1.2.1 Искусственный интеллект в компьютерном зрении в мобильных роботах .	12
1.2.2 Искусственный интеллект в компьютерном зрении в промышленных роботах	13
1.2.3 Искусственный интеллект в компьютерном зрении в промышленных роботизированных системах	14
1.3 Искусственный интеллект в компьютерном зрении в медицине.....	15
1.3.1 Диагностика патологий по изображению	16
1.3.2 Компьютерное зрение в хирургии	17
1.3.3 Вывод	18
1.4 Искусственный интеллект в обработке естественного языка	19
1.4.1 Генерация естественного языка	19
1.5 Обработка естественного языка	24
1.5.1 Статистическая обработка естественного языка	24
1.5.2 Обработка естественного языка на основе правил.....	25
1.5.3 Гибридная обработка естественного языка.....	27
1.5.4 Вывод	28
1.6 Искусственный интеллект в распознавании и синтезе речи	28

1.6.1 Искусственный интеллект в распознавании речи	29
1.6.2 Глубокое обучение и нейронные сети	29
1.6.3 Многоязычная и многодиалектная поддержка	31
1.6.4 Интеграция с другими технологиями и сервисами	32
1.6.5 Искусственный интеллект в синтезе речи.....	34
1.6.6 Вывод	39
1.7 Искусственный интеллект в рекомендательных системах и интеллектуальных системах поддержки принятия решений.....	39
1.7.1 Искусственный интеллект в рекомендательных системах.....	39
1.7.2 Искусственный интеллект в интеллектуальных системах поддержки принятия решений	49
1.7.3 Вывод	55
1.8 Генеративный искусственный интеллект.....	55
1.8.1 Генеративно–состязательные сети.....	55
1.8.2 Поточковые модели	60
1.8.3 Модели на базе трансформера.....	65
1.8.4 Вывод	70
1.9 Автоматическое машинное обучение.....	70
1.9.1 Автоматизированная генерация признаков.....	71
1.9.2 Оптимизация гиперпараметров	76
1.9.3 Обучение с подкреплением	81
1.9.4 Нейронные сети	85
1.9.5 Сильный искусственный интеллект.....	90
1.9.6 Вывод	95
1.10 Системы обработки больших объёмов данных	95
1.10.1 Обработка больших объёмов данных в режиме реального времени.....	96
1.10.2 Облачная обработка больших объёмов данных	100
1.10.3 Системы обработки больших объёмов данных с использованием машинного обучения	108

1.10.4 Вывод	113
2 Анализ существующих и перспективных рынков, развивающихся с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации и за рубежом	114
2.1 Искусственный интеллект в компьютерном зрении	114
2.2 Искусственный интеллект в обработке естественного языка	121
2.3 Искусственный интеллект в распознавании и синтезе речи	128
2.4 Искусственный интеллект в рекомендательных системах и интеллектуальных системах поддержки принятия решений	136
2.5 Генеративный искусственный интеллект	147
2.6 Автоматическое машинное обучение	154
2.7 Системы обработки больших объёмов данных	163
3 Анализ имеющихся российских и зарубежных разработок, которые оказывают (могут оказать) влияние на развитие соответствующего приоритета 20а, для их сопоставления с предложениями, представляемыми в анализируемых заявках	173
Заключение	177
Список источников	180

ВВЕДЕНИЕ

Настоящий аналитический доклад о состоянии, перспективах развития областей науки и технологий, а также рынков продукции (товаров, услуг), развивающихся с использованием результатов научной и (или) научно-технической деятельности в Российской Федерации и за рубежом подготовлен Региональным фондом научно-технического развития Санкт-Петербурга по Договору от 11.07.2023 № 6у/321180 (далее – Договор) в рамках выполнения Университетом ИТМО государственного контракта № 14.2021.244.02.0053 от 11 августа 2021 г.

В процессе работ по Договору проведен анализ:

– состояния, перспектив развития областей науки и технологий, относящихся к приоритетным направлениям 20(а) Стратегии научно-технологического развития Российской Федерации (далее – СНТР РФ): созданию систем искусственного интеллекта, в том числе машинного обучения и обработки больших объёмов данных;

– рынков продукции (товаров, услуг), развивающихся с использованием результатов научной и (или) научно-технической деятельности в Российской Федерации и за рубежом по указанным приоритетным направлениям;

– имеющихся российских и зарубежных разработок, которые оказывают (могут оказать) влияние на развитие соответствующего приоритета 20а, для их сопоставления с предложениями, представленными в анализируемых заявках, и в докладе представлены результаты указанного анализа.

В рамках анализа рассмотрены следующие области применения систем искусственного интеллекта:

- 1) компьютерное зрение;
- 2) обработка естественного языка;
- 3) распознавание и синтез речи;
- 4) рекомендательные системы и интеллектуальные системы поддержки принятия решений;
- 5) генеративный искусственный интеллект;
- 6) машинное обучение;
- 7) обработка больших объёмов данных.

В процессе анализа выявлены:

– потенциальные заказчики комплексных программ и комплексных проектов;

– имеющиеся российские и зарубежные разработки, которые оказывают (могут оказать) влияние на развитие соответствующего приоритета научно-технологического развития, для их сопоставления с предложениями, представленными в анализируемых заявках;

– информация о появлении на рынках новых инновационных решений, которые могут повлиять на актуальность ожидаемых результатов комплексных программ, комплексных проектов;

– информация о появлении новых научно-технических и(или) технологических решений и условий, не учтенных при подготовке заявок и предложений о разработке комплексной программы, комплексного проекта, способствующих сокращению сроков получения запланированного результата.

1 АНАЛИЗ СОСТОЯНИЯ И ОСНОВНЫХ ТЕНДЕНЦИЙ И ПЕРСПЕКТИВ РАЗВИТИЯ СФЕРЫ НАУКИ И ТЕХНОЛОГИЙ ПО НАПРАВЛЕНИЮ 20А СТРАТЕГИИ НАУЧНО–ТЕХНОЛОГИЧЕСКОГО РАЗВИТИЯ РОССИЙСКОЙ ФЕДЕРАЦИИ

1.1 Искусственный интеллект в компьютерном зрении

Искусственный интеллект (далее ИИ) – комплекс технологических решений, позволяющий имитировать когнитивные функции человека (включая самообучение, поиск решений без заранее заданного алгоритма и достижение инсайта) и получать при выполнении конкретных практически значимых задач обработки данных результаты, сопоставимые, как минимум, с результатами интеллектуальной деятельности человека.

Компьютерное зрение (далее КЗ) – комплекс технических средств, получающий визуальные данные, обрабатывающий и интерпретирующий их с помощью ИИ.

1.1.1 Искусственный интеллект в компьютерном зрении для детекции и идентификации объектов

Детекция и идентификация объектов – быстро растущая сфера цифровых технологий, которая направлена на решение проблем распознавания, классификации объектов, а также их расположении в пространстве.

Основными тенденциями развития искусственного интеллекта в компьютерном зрении для детекции и идентификации объектов являются:

- беспилотный транспорт,
- промышленность,
- здравоохранение.

1.1.2 Детекция и идентификация объектов в беспилотном транспорте

Беспилотный транспорт представляет собой транспортное средство, оборудованное системой автоматического управления, которое может безопасно передвигаться без участия человека. Неотъемлемой частью данного транспорта является технология детекции и идентификации объектов. Именно эта технология помогает транспорту определить какой объект находится перед ним и на каком расстоянии. Беспилотный транспорт можно классифицировать следующим образом:

- беспилотные автомобили,
- беспилотные поезда,
- беспилотные летательные аппараты,
- беспилотные суда,

— беспилотные транспортные роботы.

Научно–исследовательская и публикационная активность, связанная с направлением «Детекция и идентификация объектов в беспилотном транспорте» за последние 10 лет показана на рисунке 1.1.1 [1, 2].



Рисунок 1.1.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Детекция и идентификация объектов в беспилотном транспорте» за последние 10 лет

В работе [3] продемонстрировано усовершенствование процесса принятия решений автономными транспортными средствами с использованием теории самоопределения и нейронных сетей. Энергоэффективные нейронные сети смешанной точности используются для решения проблемы передачи сигнала на транспортные средства. В данной работе продемонстрировано увеличение точности на 80 % по сравнению с предыдущими результатами. Также разработанный алгоритм распознавания образов обеспечивает время принятия решения от 0 до 10 секунд, что увеличивает среднее время в пути без контроля оператора автономного транспортного средства на 92 %.

Для достижения идентификации объектов и обнаружения пешеходов в работе [4] предлагается механизм слияния датчиков, называемый полностью сверточными нейронными сетями для слияния лидара и камеры, который объединяет данные лидара с изображениями с нескольких камер, чтобы обеспечить оптимальное решение для обнаружения пешеходов. Модель системы предлагает отдельный алгоритм для объединения изображений при обнаружении пешеходов. Архитектура и фреймворк разработаны для полностью сверточных нейронных сетей для слияния лидара и камеры для обнаружения пешеходов. Кроме того, предлагается полнофункциональный алгоритм обнаружения и идентификации пешеходов для точного определения местоположения пешехода в диапазоне от 10 до 30 м.

В работе [5] проведен обзор глубокого обучения для автономных надводных

транспортных средств в морской среде. Как показало исследование, в отличие от смежных сфер, детекция и идентификация объектов в беспилотных судах находится на ранней стадии развития, хотя уже реализуются проекты по созданию подобных судов.

В работе [6] продемонстрирована разработка обнаружение объектов на основе глубокого обучения и восприятие сцены в плохих погодных условиях. В данной работе показано, как можно использовать модель «You Only Look Once v5» для идентификации автомобилей, светофоров и пешеходов в различных погодных условиях, что позволяет проводить идентификацию в режиме реального времени в типичной автомобильной среде. Результаты исследования показали, что предложенный подход может распознавать легковые автомобили, грузовики и другие придорожные объекты в различных погодных условиях с приемлемыми результатами, точность которых достигает 72,3 %.

1.1.3 Детекция и идентификация объектов в промышленности

Современная промышленность требует современных решений для мониторинга автоматического производства товаров и обнаружения дефектных материалов. Детекция и идентификация объектов является обязательным шагом на пути к автоматическому производству.

Научно–исследовательская и публикационная активность, связанная с направлением «Детекция и идентификация объектов в промышленности» за последние 10 лет показана на рисунке 1.1.2 [1, 2].



Рисунок 1.1.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Детекция и идентификация объектов в промышленности» за последние 10 лет

В работе [7] продемонстрирована система распознавания различных промышленных объектов, а также их дефектов с помощью технологии многоуровневого извлечения признаков

из промышленных сцен с помощью глубокого обучения. В результате была получена сверточная нейронная сеть, которая показала хорошую эффективность практически в каждом наборе данных, в тоже время среднее улучшение классификации составило практически 7 %.

Инновационное решение в области промышленной безопасности было предложено в работе [8] – метод, который направлен на обнаружение аномальных событий в больших видеоданных. Обнаружение аномалий направлено на обнаружение неожиданных или редких событий в потоках данных, например, события, которые могут причинить ущерб здоровью человека, окружающей среде, предприятию и др. Благодаря использованию двухпоточной нейронной сети с поддержкой искусственного интеллект вещей, метод имеет повышенную точность на 9,88 % по сравнению с современными методами.

В работе [9] продемонстрирован новый метод обнаружения группы объектов, расположенных под разными углами, с использованием сверточной гибридной нейронной сети на основе регионов. Исследование направлено на решение проблемы автоматизации промышленных манипуляторов, связанная с поиском более эффективного способа детекции и идентификация объектов. Результаты показывают, что точность обнаружения достигла 96,26 %.

1.1.4 Детекция и идентификация объектов в здравоохранении

Использование искусственного интеллекта для детекции и идентификации злокачественных новообразований, опухолей, бактерий и др. является важной задачей, которая упростит и ускорит процесс диагностирования широкого спектра заболеваний.

Научно–исследовательская и публикационная активность, связанная с направлением «Детекция и идентификация объектов в здравоохранении» за последние 10 лет показана на рисунке 1.1.3 [1, 2].



Рисунок 1.1.3 – Научно–исследовательская и публикационная активность, связанная с направлением «Детекция и идентификация объектов в здравоохранение» за последние 10 лет

В работах [10–14] показаны разработки, связанные с детекцией и идентификацией рака различных частей тела человека. Инновационным решением является применение искусственного интеллекта в выявлении рака легких, описанное в работе [14]. Было выявлено, что на изображениях компьютерной томографии нейронные сети дают большие надежды в обнаружении узлов в легких диаметром 3 мм. В работе [11] продемонстрирован обзор 27 систем идентификации рака желудка с помощью сверхточных нейронных сетей, точность исследований на основе данной технологии составила 77,3 – 98,7 %. Также активно развивается диагностика рака кожи, описанная в работах [10, 13]. Экспериментальные результаты показывают, что достигается превосходный процент точности (92,58 %), чувствительности (93,76 %), специфичности (91,56 %).

Активно развиваются технологии, связанные с визуализацией сердечно–сосудистой системы [15–17]. Охватываемые области визуализации на основе искусственного интеллекта варьируются от анализа изображений (например, сегментации изображения, автоматизированных измерений и, в конечном счете, автоматической диагностики) до диагностической визуализации, включая идентификацию бляшек, оценку уязвимости бляшек, оценку гемодинамики миокарда, такую как компьютерная томография на основе глубокого обучения, и оценку прогноза риска. В частности, способность алгоритмов искусственного интеллекта ставить более точные диагнозы полезна врачам для выявления заболеваний на ранних стадиях их течения, чтобы спланировать правильное лечение.

В работах [18–21] продемонстрированы перспективные разработки по борьбе с COVID–19 с помощью детекции и идентификации на основе искусственного интеллекта. В работе [21] была обучена сверточная нейронная сеть на рентгенограмме туловища, способная обеспечить точность определения COVID–19 до 95 %. Другой исследовательской группой [19] были получены схожие результаты, но с большей точностью – 98 %.

Инновационное решение борьбы с COVID–19 было описано в работе [18].

1.2 Искусственный интеллект в компьютерном зрении в робототехнике

В настоящее время компьютерное зрение является одной из главных составляющей робототехники. Мобильные роботы, промышленные роботы, манипуляторы, промышленные роботизированные системы хорошо выполняют повторяющие рутинные задачи, но когда возникает аномальное событие, которое не заложено в тексте программы, данные устройства практически беспомощны. Именно искусственный интеллект в компьютерном зрении помогает решить данную проблему.

Основными тенденциями развития искусственного интеллекта в компьютерном зрении в робототехнике являются:

- мобильные роботы,
- промышленные роботы,
- промышленные роботизированные системы.

1.2.1 Искусственный интеллект в компьютерном зрении в мобильных роботах

Мобильные роботы с компьютерным зрением – это автономные устройства, которые оснащены камерами и компьютерными алгоритмами обработки изображений. Они могут использоваться для различных задач, таких как автоматическое сбор данных, мониторинг окружающей среды, обнаружение и распознавание объектов и т.д. Благодаря компьютерному зрению, мобильные роботы могут с легкостью выполнять сложные задачи, которые ранее требовали наличия человека.

Научно–исследовательская и публикационная активность, связанная с направлением «Искусственного интеллекта в компьютерном зрении в мобильных роботах» за последние 10 лет показана на рисунке 1.2.1 [1, 2].



Рисунок 1.2.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Искусственного интеллекта в компьютерном зрении в мобильных роботах» за последние 10 лет

В работе [22] продемонстрирован навигационный подход на основе нейронных сетей для автономных мобильных роботов. Производительность метода принятия решений на основе нейронных сетей была проверена с использованием данных обучения и тестирования, чтобы показать уровень точности, близкий к 90 %.

Подход к сегментации сложных поверхностей на основе изображений с использованием инструментов машинного обучения (далее МО) во время движения мобильных роботов был продемонстрирован в работе [23]. Обучающий набор данных включал сцены следующих типов: сцены с лестницами, сцены с ровными поверхностями,

сцены с изолированными препятствиями и с группами препятствий. Согласно результатам тестирования, разработанный подход демонстрирует приличное качество сегментации на изображениях с ровными поверхностями 90,2 % и лестницами 71,3 %, а также сохраняет некоторую устойчивость к изменениям уровней освещенности сцены.

В работе [24] продемонстрирована разработка технология для навигации мобильных роботов с помощью монокулярного зрения с использованием алгоритма внутреннего анализа, который объединяет различные уровни сверточной нейронной сети с процессом принятия решений иерархическим образом. Как показали эксперименты, точность такой технологии достигает 77 %.

1.2.2 Искусственный интеллект в компьютерном зрении в промышленных роботах

Промышленный робот – автоматически управляемый, перепрограммируемый манипулятор, программируемый по трем или более степеням подвижности, который может быть установлен стационарно или на мобильной платформе для применения в целях промышленной автоматизации.

Манипулятор – устройство, механизм которого обычно состоит из нескольких звеньев, вращающихся или перемещающихся поступательно друг относительно друга с целью взятия и/или перемещения объектов (деталей или инструмента), как правило, по нескольким степеням свободы.

Научно–исследовательская и публикационная активность, связанная с направлением «Искусственного интеллекта в компьютерном зрении в промышленных роботах» за последние 10 лет показана на рисунке 1.2.2 [1, 2].



Рисунок 1.2.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Искусственного интеллекта в компьютерном зрении в промышленных роботах» за последние 10 лет

В работе [25] продемонстрирован новый подход к оперативной калибровке промышленных роботов с помощью системы захвата движения и нейронной сети «Extreme Learning Machine». Комбинация методов идентификации параметров на основе модели и компенсации на основе нейронной сети является эффективным решением для исправления всех источников ошибок робота. После калибровки точность определения положения робота составляет от 7,440 до 0,159 мм, а точность ориентации – от 3,073 до 0,077 градусов.

Метод автоматического обучения глубокой нейронной сети для технического зрения промышленного робота показан в работе [26]. Результаты показывают, что временные затраты на автоматическое аннотирование данных с помощью алгоритма «Single-Shot-Detector» на 77,75 % меньше, чем при ручном аннотировании данных, а скорость распознавания модели, обученной с использованием данных автоматической маркировки, составляет 97,75 %, что подтверждает эффективность и преимущества нового метода.

В работе [27] разработан метод обнаружения положения сложенных объектов для захвата промышленным роботом с помощью сверточной нейронной сети и нескольких автономных датчиков. В среде с разбросанными и сложенными объектами результаты эксперимента показывают, что метод, основанный на сети «ResNet», превосходит другие три сети, а средний показатель успешности обнаружения позы захвата на основе сетей «ResNet», «InceptionV2», «VGG» и «LetNet» составляет 90,67 %, 82,67 %, 86,67 % и 87,33 % соответственно, что подтверждает эффективность и преимущества разработанного метода.

1.2.3 Искусственный интеллект в компьютерном зрении в промышленных роботизированных системах

Промышленная роботизированная система – система, состоящая из промышленного робота, конечных эффекторов и любых машин, оборудования, устройств, внешних вспомогательных осей или датчиков, поддерживающих выполнение роботом своей задачи. Они работают непрерывно через повторяющиеся циклы движения в соответствии с инструкциями набора команд, называемых программой.

Научно–исследовательская и публикационная активность, связанная с направлением «Искусственного интеллекта в компьютерном зрении в промышленных роботизированных системах» за последние 10 лет показана на рисунке 1.2.3 [1, 2].



Рисунок 1.2.3– Научно–исследовательская и публикационная активность, связанная с направлением «Искусственного интеллекта в компьютерном зрении в промышленных роботизированных системах» за последние 10 лет

Разработка системы оценки положения на основе зрения для роботизированной обработки и повышение её точности с использованием нейронных сетей «LSTM» и разреженной регрессии показана в работе [28]. Предложенные методы подтверждены экспериментальным исследованием, выполненным с использованием робота «KUKA KR240 R2900 ultra» при обработке детали «NAS 979». Достижимые абсолютные ошибки положения составили в среднем 5,47 мм, 2,9 мм и 2,05 мм для обработки «NAS 979» и в среднем 5,35 мм, 2,17 мм и 0,86 мм для произвольной формы.

В работе [29] продемонстрирована разработка нового гибридного автономного метода программирования для роботизированных сварочных систем. В данной работе объединяется система автоматизированного проектирования, компьютерное зрение «vision-based», «vision & CAD interactive», чтобы преодолеть ограничения современных методов автоматической генерации программ. Как заявляют авторы, этот метод может значительно повысить эффективность, точность и гибкость роботизированной сварочной системы. Эксперименты показали, что отклонение в начальных точках сварочных швов уменьшилась до 0,7 мм.

В работе [30] показана разработка для обнаружения объектов с использованием рандомизации домена «Sim2Real» для приложений роботизированных систем. Авторы смогли сократить приблизиться к работе в реальном времени, достигнув 86,32 % и 97,38 % баллов по mAP₅₀. Данное решение разрабатывалась для промышленного использования, процесс генерации данных занимает менее 0,5 с на изображение, а обучение длится всего 12 ч на графическом процессоре «GeForce RTX 2080 Ti».

1.3 Искусственный интеллект в компьютерном зрении в медицине

В настоящее время существует проблемы диагностики патологий по изображению

(фото, МРТ, КТ, УЗИ и др.), в реальном времени во время операций, связанная с человеческим фактором. Искусственный интеллект в компьютерном зрении в медицине является решением данной проблемы.

Основными тенденциями развития искусственного интеллекта в компьютерном зрении в медицине являются:

- диагностика патологий по изображению,
- компьютерное зрение в хирургии.

1.3.1 Диагностика патологий по изображению

Искусственный интеллект в компьютерном зрении в медицине для диагностики патологий по изображению применяется в:

- компьютерной томографии,
- ультразвуковом исследовании,
- магнитно–резонансной томографии,
- оптической когерентной томографии,
- эндоскопии.

Научно–исследовательская и публикационная активность, связанная с направлением «Диагностика патологий по изображению» за последние 10 лет показана на рисунке 1.3.1 [1, 2].



Рисунок 1.3.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Диагностика патологий по изображению» за последние 10 лет

В работе [31] представлены разработки в области искусственного интеллекта для анализа интерстициальных заболеваний легких на компьютерной томографии грудной клетки. Уже на данный момент достигнута точность в 78–91 % при определении интерстициальных заболеваний легких.

В работе [32] был разработан метод диагностики рака молочной железы на основе искусственного интеллекта с использованием ультразвуковых изображений и генератора глубоких признаков на основе сетки. Разработанная модель классификации изображений на основе генерации глубоких признаков на основе сетки достигла точности классификации 97,18 % на ультразвуковых изображениях для трех классов, а именно злокачественных, доброкачественных и нормальных.

Метод классификации опухолей головного мозга на основе гибридного оптимизированного анализа по нескольким признакам с использованием набора данных магнитно-резонансной томографии продемонстрирован в работе [33]. Максимальная эффективность классификации опухолей головного мозга составила 98,8 %. Компоненты и производительность предлагаемой структуры показывают, что это новая и надежная структура классификации.

В работе [34] была продемонстрирована разработка оптической когерентной томографии для выявления заболеваний глаз на основе изображений с использованием глубокой сверточной нейронной сети. Данная модель эффективно классифицирует различные нарушения зрения, а именно неоваскуляризацию хориоидеи, диабетический макулярный отек, друзы на изображениях оптической когерентной томографии с точностью 95,6 – 97,79 %.

В работе [35] был разработан искусственный интеллект, способный в режиме реального времени обнаружить очаговые поражения и диагностировать новообразования желудка с помощью эндоскопии в белом свете. В исследование были включены данные более 10 000 пациентов, которым проводилась эндоскопия верхних отделов. Чувствительность составила 96,9 % и 95,6 % для обнаружения поражений желудка и 92,9 % и 91,7 % для диагностики новообразований у внутренних и внешних пациентов соответственно.

1.3.2 Компьютерное зрение в хирургии

В последние годы искусственный интеллект показал большие перспективы в медицине. Однако проблемы с объяснимостью затрудняют применение ИИ в клинических условиях. Были проведено множество исследований в области объяснимого искусственного интеллекта, чтобы преодолеть ограничения, связанные с природой черного ящика методов искусственного интеллекта.

Научно-исследовательская и публикационная активность, связанная с направлением «Компьютерное зрение в хирургии» за последние 10 лет показана на рисунке 1.3.2 [1, 2].



Рисунок 1.3.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Компьютерное зрение в хирургии» за последние 10 лет

В работе [36] был разработан искусственный интеллект для автоматической оценки критического взгляда на безопасность при лапароскопической холецистэктомии с использованием глубокого обучения. Точность разработанной модели достигла 71,9 %.

В работе [37] проведен анализ методов самостоятельного обучения для хирургического компьютерного зрения. Было выявлено, что при правильном переносе методов самоконтролируемого обучения «MoCo v2», «SimCLR», «DINO», «SwAV» в хирургию, как описано в работе, можно добиться существенного увеличения производительности по сравнению с обычным использованием метода самоконтролируемого обучения – до 7,4 % при распознавании фазы и 20 % при обнаружении присутствия инструмента, – а также к современным подходам распознавания фазы с полуконтролем до 14 %.

Метод обнаружения пятен крови с использованием алгоритмов, основанных на компьютерном зрении, и их связь с послеоперационными результатами при торакоскопических лобэктомиях был разработан в работе [38]. Пиксели крови в хирургических видеороликах были идентифицированы с помощью алгоритмов компьютерного зрения. Доля пикселей крови была рассчитана для ключевых видеокadres, чтобы обобщить информацию о пятнах крови во время операции. Данный метод был разработан для эффективного выявления из хирургических видеороликов информацию, которая обладает хорошей способностью предсказывать послеоперационные результаты с точностью до 95 %.

1.3.3 Вывод

Основными тенденциями развития искусственного интеллекта в компьютерном зрении являются внедрения и разработка: детекции и идентификации объектов для беспилотного транспорта, детекция и идентификация объектов в промышленности, детекция и

идентификация объектов в здравоохранении, компьютерное зрение в хирургии, в диагностировании патологий по изображению, компьютерное зрение в мобильных и промышленных роботах.

1.4 Искусственный интеллект в обработке естественного языка

Искусственный интеллект в обработке естественного языка – раздел искусственного интеллекта, который направлен на взаимодействие между компьютерами и человеческим языком, в частности на то, как программировать компьютеры для обработки и анализа больших объёмов данных на естественном языке.

1.4.1 Генерация естественного языка

Генерация естественного языка – это использование программирования искусственного интеллекта для создания письменных или устных текстов из набора данных.

Основными тенденциями развития искусственного интеллекта в генерации естественного языка являются:

- здравоохранение,
- розничная торговля, банковское дело, финансовые услуги и страхование,
- реклама и средства массовой информации.

1.4.1.1 Генерация естественного языка в здравоохранении

Генерация естественного языка играет важную роль в медицине, где точность и ясность информации имеют ключевое значение. Некоторые из областей, где генерация естественного языка может быть использована в медицине, включают создание отчётов о пациентах, подготовку медицинских статей, автоматизацию отчётности о заболеваниях, изложение медицинских терминов и текстов простым языком и многое другое.

Научно–исследовательская и публикационная активность, связанная с направлением «Генерация естественного языка в здравоохранении» за последние 10 лет показана на рисунке 1.4.1 [1, 2].



Рисунок 1.4.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Генерация естественного языка в здравоохранении» за последние 10 лет

В работе [39] представлена модель порождающего языка «BioBART», которая адаптирует модель генерации естественного языка «BART» к биомедицинской области. Были сопоставлены различные задачи генерации биомедицинского языка, включая диалог, обобщение, связывание объектов и распознавание именованных объектов. Предварительное обучение «BioBART» на «PubMed abstracts» повысило производительность по сравнению с BART и установило надежные исходные условия для нескольких задач. Кроме того, были проведены абляционные исследования задач предварительного обучения для «BioBART» и обнаружено, что перестановка предложений оказывает негативное влияние на последующие задачи.

В работе [40] была представлена иерархическая структура для комплексного паралингвистического анализа речи, охватывающая пол, эмоции и распознавание лжи. Было исследовано взаимодействие между различными паралингвистическими явлениями, используя гендерную информацию для прогнозирования эмоциональных состояний, а результат распознавания эмоций применялся для прогнозирования правдивости речи. Для проведения экспериментов использовались несколько наборов данных, включая aGender, Ruslana, EmoDB и DSD. Эксперименты показали, что гендерно-специфические модели повышают эффективность автоматического распознавания речевых эмоций и обнаружения лжи. Полученные результаты перекрестной проверки для обнаружения обмана превосходят существующий уровень техники на абсолютные 2,8 %. В целом, разработанная модель с успехом демонстрирует возможности применения паралингвистического анализа речи для решения различных задач в области общения и коммуникации.

При поиске информации, не отраженной в удобных для пациентов документах, потребители медицинских услуг могут обратиться к исследовательской литературе. Однако чтение медицинских статей может быть сложным опытом. Чтобы улучшить доступ к

медицинским документам в работе [41] представлен новый интерактивный интерфейс «Paper Plain» с четырьмя функциями на основе генерации естественного языка: определения незнакомых терминов, краткое изложение объёмных текстов на простом языке, набор ключевых вопросов, которые помогут читателям задать вопрос, краткое изложение ответов простым языком. Результаты исследования показывают, что пользователи могут получить информацию из медицинских статей в кратком виде на доступном для них языке.

В работе [42] продемонстрирована разработка генерации медицинских отчетов. Для решения данной задачи была разработана технология «ASGK» (Auxiliary Signal–Guided Knowledge Encoder–Decoder). Экспериментальные результаты подтверждают, что технология «ASGK» может превзойти предыдущие подходы как в классификации медицинской терминологии, так и в показателях генерации абзацев.

Также генерацию естественного языка предлагается использовать для изучения психоза и психологических расстройств. В работе [43] предлагают использовать системы генерации естественного языка, в качестве инструментов для изучения лингвистических особенностей психологических расстройств. В работе утверждается, что данное направление исследований позволит лучше понять взаимосвязь между языком и психозом и потенциально проложит путь к новым терапевтическим подходам для решения этой проблемы.

1.4.1.2 Генерация естественного языка в розничной торговле, банковском деле, финансовых услугах и страховании

Генерация естественного языка является важным инструментом в розничной торговле, банковском деле, финансовых услугах и страховании, который позволит: создавать краткое описание товаров и услуг, генерировать и анализировать отчеты о финансовых операциях, генерировать ответы для голосовых помощников, а также чат–ботов.

Научно–исследовательская и публикационная активность, связанная с направлением «Генерация естественного языка в розничной торговле, банковском деле, финансовых услугах и страховании» за последние 10 лет показана на рисунке 1.4.2 [1, 2].



Рисунок 1.4.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Генерация естественного языка в розничной торговле, банковском деле, финансовых услугах и страховании» за последние 10 лет

В работе [44] продемонстрирована разработка генерации доказательств на естественном языке с помощью поиска, управляемого верификатором, которая в перспективе может внедрена в судебной практике. В этой работе фокусируются на генерации доказательств: учитывая гипотезу и набор подтверждающих фактов, модель генерирует дерево доказательств, указывающее, как вывести гипотезу из подтверждающих фактов. По сравнению с генерацией всего доказательства за один раз, пошаговая генерация может лучше использовать композиционность и обобщать на более длинные доказательства. Данный метод имеет точность 33,3 %.

В работе [45] рассмотрены перспективы внедрения генерации финансовых отчетов. Аналитики столкнулись с проблемами недостоверности и разрозненности информации. Главной причиной этого являлось несоответствии между темой входных новостей и выходными знаниями в целевых отчетах. Несмотря на данные проблемы, как показано в работе, производительность генератора финансовых отчетов достигает 48 %.

Актуальной является разработка [46] генерации информации о страховании для эффективного управления рисками с помощью машинного обучения. Сгенерированные данные могут оказаться полезными для принятия решений и разработки стратегии в областях и задачах бизнеса. Разработанный метод машинного обучения, основанный на наборе выбранных функций, предложенных алгоритмами выбора функций, превзошли полный набор функций для набора реальных страховых наборов данных. В частности, 20 % и 50 % функций в пяти наборах данных улучшили последующую кластеризацию и производительность классификации по сравнению с целыми наборами данных. Это указывает на потенциал выбора функций в страховом секторе как для повышения производительности модели, так и для

выделения важных функций для анализа бизнеса.

1.4.1.3 Генерация естественного языка в рекламе и средствах массовой информации

Генерация естественного языка в рекламе и средствах массовой информации – это процесс создания текстов, которые звучат естественно и привлекают внимание аудитории. Для успешной генерации естественного языка в рекламе и средствах массовой информации необходимо учитывать целевую аудиторию, ее интересы и предпочтения, а также использовать правильную грамматику и стиль написания текстов.

Научно–исследовательская и публикационная активность, связанная с направлением «Генерация естественного языка в рекламе и средствах массовой информации» за последние 10 лет показана на рисунке 1.4.3 [1, 2].



Рисунок 1.4.3 – Научно–исследовательская и публикационная активность, связанная с направлением «Генерация естественного языка в рекламе и средствах массовой информации» за последние 10 лет

В работе [47] показана разработка для поддержки контент–маркетинга с помощью генерации естественного языка. Разработанная технология способна не только сократить расходы, связанные созданием контент–маркетинга, но и частично заменить специалистов в этой области. Как показывают результаты, сгенерированные тексты практически невозможно отличить от текстов, написанных специалистами в данной области.

В работе [48] продемонстрирована разработка управляемой генерации текста с использованием предварительно обученных языковых моделей на основе «Transformer». Данная технология направлена на генерацию разнообразного и беглого текста, что позволяет учитывать интересы и предпочтения аудитории, а также гибко изменять стиль написания текстов.

В работе [49] продемонстрирована разработка генератора естественного языка для копирайтинга продуктов и опыт её внедрения. Разработка направлена на привлечение интереса пользователей и улучшение пользовательского опыта путем выделения характеристик продукта с помощью текстовых описаний. К сентябрю 2021 года было сгенерировано 2,53 миллиона описаний продуктов и улучшил общий усредненный показатель переходов и коэффициент конверсии на 4,22 % и 3,61 % по сравнению с базовыми показателями, соответственно, в годовом исчислении. Совокупный валовой объём товаров, созданный разработанной системой, улучшен на 213,42 % по сравнению с показателем на февраль 2021 года, что наглядно показывает эффективность внедрения генератора естественного языка.

1.5 Обработка естественного языка

Обработке естественного языка (далее NLP, от англ. – Natural language processing) – это анализ текста и речи на естественном языке, а также языка жестов (знаков) для получения информации, подлежащей обработке средствами вычислительной техники.

Основными тенденциями развития искусственного интеллекта в обработке естественного языка являются:

- статистическая обработка естественного языка,
- обработка естественного языка на основе правил,
- гибридная обработка естественного языка.

1.5.1 Статистическая обработка естественного языка

Статическая обработка естественного языка – это процесс предсказания следующего слова в последовательности с учётом слов, которые ему предшествуют. Преимущество этого метода состоит в том, что он способен обрабатывать большие объёмы текстовых данных и неструктурированные данные. Статистическое моделирование помогает: давать прогнозы и оценивать риски на основе текста, обобщать текст, делать выводы и др.

Научно–исследовательская и публикационная активность, связанная с направлением «Статистическая обработка естественного языка» за последние 10 лет показана на рисунке 1.5.1 [1, 2].



Рисунок 1.5.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Статистическая обработка естественного языка» за последние 10 лет

В работе [50] продемонстрирован ряд новых моделей обработки естественного языка для применений в маркетинге. В результате было выяснено, что модели ANN (Artificial Neural Network), NB (Naïve Bayes) и RF (Random Forest) подходят для точного прогнозирования и оценки рисков, когда KNN (K–Nearest Neighbor) и DIST (Decision Trees) имеют преимущества в интерпретируемости. Также в работе приведена статистика, из которой видно, что внедрение обработки естественного языка способно с точностью до 95 % предсказывать настроение клиентов.

В работе [51] разработан метод автоматизированной маркировки ошибок в радиационной онкологии с помощью статистической обработки естественного языка. Все полученные модели, кроме одной, имели отличную производительность, выраженную в эффективности работы достигающей 90 % и точности в 100 % случаев.

В работе [52] разработан метод обнаружения настроения с использованием обработки естественного языка и машинного обучения в социальных взаимодействиях. В работе используются крупномасштабные наборы данных из социальных сетей для выполнения анализа настроений с помощью алгоритмов машинного обучения и модели глубокого обучения «BERT» для классификации настроений. Разработанная модель может определить настроение сообщения с точностью до 97,78 %.

1.5.2 Обработка естественного языка на основе правил

Обработка естественного языка на основе правил использует predetermined набор правил для анализа и обработки текстовых данных. Этот подход является одним из самых старых в области NLP и до сих пор активно используется, и модернизируется. Данный метод проверен временем и является самым стабильным. Этот подход является быть более точным и

предсказуемым, чем статистический подход. Он также может обрабатывать малые объёмы данных, такие как запросы к поисковым системам или команды голосовых помощников.

Научно–исследовательская и публикационная активность, связанная с направлением «Обработка естественного языка на основе правил» за последние 10 лет показана на рисунке 1.5.2 [1, 2].



Рисунок 1.5.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Обработка естественного языка на основе правил» за последние 10 лет

Обработка естественного языка на основе правил нашла применение в медицине. В работах [53–55] продемонстрированы перспективные технологии обработки естественного языка на основе правил в области здравоохранения.

В работе [55] рассмотрен и оценён метод обработки естественного языка для определения расширенного показателя по шкале инвалидности из электронных медицинских карт пациентов с рассеянным склерозом. Модель обучена на 17 452 электронных медицинских картах 4 906 пациентов с рассеянным склерозом. Как результат комбинированная модель обработки естественного языка может извлекать и точно прогнозировать оценки EDSS из записей пациентов с точностью до 90 %. Этот подход может быть автоматизирован для эффективного извлечения информации в клинических условиях и исследовательских учреждениях.

В работе [54] была разработана система обработки естественного языка на основе правил для автоматизации извлечения данных об инсульте. Как показывают результаты, разработанная система имеет хорошую производительность для извлечения атрибутов, связанных с инсультом, из отчетов рентгенологии в произвольном виде, когда применяется к внешнему набору данных.

Также важной технологией является алгоритм обработки естественного языка для идентификации младенцев с лихорадкой [53]. Модель была обучена на 8 935 выписках и

записях о приеме. Как показывают результаты, чувствительность метода достигает 99 %.

В работе [56] показана обработка ошибок на основе нейронной сети при обработке естественного языка. Разработанный метод включает гибридный подход к разработке средства проверки грамматики тамильского языка, поскольку тамильский язык обладает множеством грамматических особенностей. Из всех средств проверки ошибок частота обнаружения ошибок для обработчика ошибок с согласными составляет 94 %, что является самым высоким показателем из аналогичных технологий.

В работе [57] продемонстрирована разработка для понимания контекста при извлечении и геокодировании исторических наводнений, штормов и мер адаптации. Разработанная модель способна извлекать такую информацию, как перекрытия улиц, затраты на проект и показатели. В результате проекта было обнаружено 435 353 мест, где за последние два десятилетия происходили штормы по всей территории Соединенных Штатов.

1.5.3 Гибридная обработка естественного языка

Гибридная обработка естественного языка – это метод, который сочетает в себе преимущества и возможности статистической обработки и обработки на основе правил. В этом методе используются как статистические модели, основанные на анализе больших объёмов текстовых данных, так и правила, основанные на лингвистических знаниях. Преимуществом данного метода является возможность использовать правила для уточнения результатов статистических методов и наоборот. Гибридный подход также может быть более эффективным при обработке сложных и неоднозначных текстовых данных.

Научно–исследовательская и публикационная активность, связанная с направлением «Гибридная обработка естественного языка» за последние 10 лет показана на рисунке 1.5.3 [1, 2].



Рисунок 1.5.3 – Научно–исследовательская и публикационная активность, связанная с

направлением «Гибридная обработка естественного языка» за последние 10 лет

В работе [58] продемонстрирован метод глубокого обучения с использованием многозадачного обучения в понимании естественного языка. Как показало исследование, языковые модели, основанные на неконтролируемом методе обучения, демонстрируют потенциальное улучшение производительности. Однако они сталкиваются с проблемой разработки универсальной основы для языковой модели, которая улучшит производительность многозадачной эффективности обработки естественного языка и обобщенного представления знаний. Объединение представленных подходов может привести к значительному повышению эффективности обработки естественного языка.

Также гибридная обработка естественного языка активно внедряется в робототехнику. Так, в работе [59] приведены два экспериментальных теста с использованием колаборативных роботов, естественного языка и метода взаимодействия на основе блоков. Полученные результаты показывают, что данный подход является не только эффективным, но делает задачу программирования простой и эффективной для нетехнических пользователей.

В работе [60] разработана модель, основанная на глубоком обучении, использующая гибридный подход к извлечению признаков для анализа потребительских настроений. Возможными поставщиками являются встроенные приложения для обмена сообщениями, такие как Telegram и WhatsApp. Для эффективного извлечения признаков был введен гибридный метод, включающий функции, связанные с обзором, и функции, связанные с аспектами, для построения отличительного гибридного вектора признаков, соответствующего каждому обзору. Классификация настроений выполняется с использованием классификатора глубокого обучения «LSTM». Модель обеспечивает среднюю точность, средний отзыв и средний балл F1 в 94,46 %, 91,63 % и 92,81 % соответственно.

1.5.4 Вывод

Таким образом, основными тенденциями развития искусственного интеллекта в обработке естественного языка являются внедрения и разработка: генерация естественного языка в здравоохранении, статистическая обработка естественного языка, гибридная обработка естественного языка, обработка естественного языка на основе правил.

1.6 Искусственный интеллект в распознавании и синтезе речи

Синтезатор речи – это техническое средство, преобразующее текст, составленный на различных языках, в звуковой сигнал, воспринимаемый слушателем как аналог человеческой речи при различных степенях разборчивости и естественности звучания.

Система распознавания речи – система, не требующая обучения на речи конкретного пользователя, пригодная для любого пользователя из выбранной группы (носителей языка,

взрослых и т.д.).

1.6.1 Искусственный интеллект в распознавании речи

Искусственный интеллект в распознавании речи – это технология и область исследования, направленная на анализ и преобразование человеческой речи в письменный текст или в понятные для компьютера инструкции. Это достигается с помощью алгоритмов машинного обучения, которые анализируют звуки, интонации и акценты речи, чтобы транскрибировать и интерпретировать их.

Распознавание речи используется во многих приложениях, таких как голосовые помощники (например, Siri, Alexa, Алиса и Google Assistant), автоматическая транскрипция аудио, системы автоматического перевода, устройства для глухих и слабослышащих, а также в системах, предназначенных для управления устройствами с помощью голосовых команд.

Основными тенденциями развития искусственного интеллекта в распознавании речи являются:

- глубокое обучение и нейронные сети,
- многоязычная и многодиалектная поддержка,
- интеграция с другими технологиями и сервисами.

1.6.2 Глубокое обучение и нейронные сети

Для распознавания речи глубокое обучение и нейронные сети используются в качестве методов машинного обучения. Обычно данные подаются на вход нейросети в формате аудиосигнала, который затем преобразуется в спектрограмму. Благодаря использованию глубокого обучения и нейронных сетей, возможно достичь достаточно высокой точности распознавания речи, что приводит к созданию новых возможностей для автоматизации бизнес-процессов и улучшения пользовательского опыта в приложениях и устройствах.

Научно-исследовательская и публикационная активность, связанная с направлением «Глубокое обучение и нейронные сети» за последние 10 лет показана на рисунке 1.6.1 [1, 2].



Рисунок 1.6.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Глубокое обучение и нейронные сети» за последние 10 лет

В работе [61] продемонстрирована разработка новой системы распознавания речи для распознавания тональных речевых сигналов с использованием сверточной нейронной сети. Метод идентификации тональной речи и добавления инструментальных знаний работает лучше, чем существующие и традиционные подходы. Результаты эксперимента демонстрируют значительную производительность существующей архитектуры расписывания речи, обеспечивая уровень точности 89,15 % для непрерывных речевых сигналов с различными тонами.

В настоящее время активно ведутся разработки с использованием рекуррентной нейронной сети. Например, в работе [62] разработана система автоматического распознавания речи на основе рекуррентной нейронной сети с длительной кратковременной памятью (далее LSTM, от англ. Long short-term memory) Данная модель превосходит традиционные подходы к машинному обучению, например, искусственная нейронная сеть. В предлагаемой модели рекуррентная нейронная сеть включена в качестве “элемента забывания” в блок памяти, чтобы обеспечить сброс состояний ячейки в начале подпоследовательностей. Это позволяет системе эффективно обрабатывать непрерывные входные потоки без обязательного увеличения требуемой полосы пропускания. В предлагаемой модели стандартная архитектура сети долговременная кратковременная память модифицирована для эффективного использования параметров модели. Разработанная модель превзошла другие модели глубокого обучения с точностью 99,36 % в хорошо зарекомендовавшем себя общедоступном тестовом наборе данных «spoken English digit dataset».

Также активно ведутся разработки в области распознавание эмоциональной речи. В

работе [63] разработан метод распознавание эмоциональной речи с использованием глубоких нейронных сетей. Параметры признаков, используемые для распознавания, включают спектральные коэффициенты Mel и другие параметры, связанные со спектром и интенсивностью речевого сигнала. Увеличение данных использовалось путем изменения голоса и добавления белого шума. Результаты показывают, что модель «GRU» дала самую высокую среднюю точность распознавания – 97,47 %. Этот результат превосходит существующие исследования по распознаванию эмоций в речи с помощью «IEMOCAP corpus».

1.6.3 Многоязычная и многодиалектная поддержка

Многодиалектная поддержка в распознавании речи – это возможность системы распознавания речи работать с различными диалектами одного и того же языка. Например, в английском языке есть британский, американский, австралийский и другие диалекты, и система распознавания речи с многодиалектной поддержкой сможет распознавать и понимать речь на любом из них.

Многоязычная поддержка в распознавании речи – это возможность системы распознавания речи работать с различными языками, то есть распознавать и понимать речь на нескольких языках одновременно. Многоязычные модели автоматического распознавания речи показали большие перспективы в последние годы из-за простого процесса обучения модели и развертывания.

Научно–исследовательская и публикационная активность, связанная с направлением «Многоязычная и многодиалектная поддержка» за последние 10 лет показана на рисунке 1.6.2 [1, 2].



Рисунок 1.6.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Многоязычная и многодиалектная поддержка» за последние 10 лет

В работе [64] разработана новая настраиваемая многоязычная модель (далее СММ, от англ. configurable multilingual model). Разрабатываемая СММ может быть развернута в любом пользовательском сценарии, где можно предварительно выбрать любую комбинацию языков. После 75 тысяч часов расшифровки анонимизированных многоязычных данных Microsoft и оценки с помощью наборов тестов на 10 языках, были получены результаты, из которых ясно, что предлагаемая СММ превосходит универсальную многоязычную модель на 26,0 %, 16,9 % и 10,4 % относительного сокращения ошибок в словах при выборе пользователем 1, 2 или 3 языков соответственно.

В работе [65] разработан метод для увеличения текстовых данных для распознавания речи с переключением кодов на арабском и английском языках. Данный метод основан на случайных лексических заменах и ограничении эквивалентности, используя выровненные пары переводов для создания случайного и грамматически корректного контента. Эмпирические результаты показывают относительное снижение сложности языковой модели на 65,5 % и 7,7 % ASR WER на двух наборах тестов.

В работе [66] разработан метод многозадачного обучения с помощью вспомогательного преобразователя перекрестного внимания для распознавания речи на нескольких диалектах с низким уровнем ресурсов. В модели есть два потока задач: основной поток, который распознает речь, и вспомогательный поток, определяющий диалект. Экспериментальные результаты по задаче распознавания тибетской многодиалектной речи показывают, что модель превосходит однодиалектную модель и многодиалектную модель, основанную на жестком совместном использовании параметров, за счет снижения средней частоты ошибок в слогах на 30,22 % и 3,89 % соответственно.

1.6.4 Интеграция с другими технологиями и сервисами

Искусственный интеллект в распознавании речи может быть интегрирован с различными технологиями, чтобы обеспечить более эффективную работу системы. Например, система распознавания речи может быть интегрирована с облачными вычислениями, чтобы усилить производительность и сэкономить время на обработке данных. Также интеграция с другими технологиями, машинным обучением и глубоким обучением, может помочь системе стать более точной и детализированной в своих предсказаниях.

Использование искусственного интеллекта в распознавании речи также может быть интегрировано в другие технологии – робототехника, автоматизация и Интернет вещей, чтобы создать более умные и автоматизированные системы.

Научно–исследовательская и публикационная активность, связанная с направлением

«Интеграция с другими технологиями и сервисами» за последние 10 лет показана на рисунке 1.6.3 [1, 2].



Рисунок 1.6.3 – Научно–исследовательская и публикационная активность, связанная с направлением «Интеграция с другими технологиями и сервисами» за последние 10 лет

В работе [67] показан пример интеграции синтеза речи и визуального распознавания речи с использованием LSTM и глубокой сверточной нейронной сети. Аудиовизуальное распознавание речи является одной из новых областей исследований и решает проблему искажения звуком. Точность метода составила 90 % для аудио–распознавания речи, 71 % для визуального распознавания речи и 91 % для аудиовизуального распознавания речи.

Исследования по распознаванию эмоций в последние годы демонстрируют растущую тенденцию из–за повышения уровня стресса у людей. В работе [68] продемонстрирован метод глубокого слияния классификаторов для распознавания эмоций с использованием речи и статических изображений. Интегрированная платформа с использованием глубокого слияния классификаторов демонстрирует точность 94,26 %.

В работе [69] продемонстрирована разработка терминала самообслуживания на основе синтеза речи. Точность распознавания речи такого терминала составляет приблизительно 80 %.

В работе [70] показана разработка, обеспечивающая взаимосвязь между роботом и человеком на основе компьютерного зрения и распознавания речи. Основой метода является расширение формата описания голосовых команд, ориентированных на выполнение задач, а также его интеграция с гибкими редактируемыми шаблонами контуров, используемыми для классификации контуров, полученных из систем распознавания изображений. Данная работа является направлена на разработку индивидуальных приложений голосового управления на основе зрения для промышленных роботов.

1.6.5 Искусственный интеллект в синтезе речи

Искусственный интеллект в синтезе речи – это технология и область исследования, направленная на преобразование текстовой информации в аудио речь, имитирующую человеческий голос. Это достигается с помощью алгоритмов машинного обучения, которые анализируют текст и синтезируют его в аудио сигнал, используя знания о фонетике, интонации, акцентах и других аспектах человеческой речи.

Синтез речи широко используется в различных приложениях, таких как голосовые помощники, аудиокниги, системы чтения для слабовидящих и незрячих, автоматические информационные системы (например, GPS–навигация), робототехника и многое другое. Некоторые из наиболее известных систем синтеза речи включают Yandex SpeechKit, Google Text-to-Speech, Amazon Polly и Microsoft's Cortana.

Основными тенденциями развития искусственного интеллекта в синтезе речи являются:

- естественность и выразительность в синтезе речи,
- пользовательский голос и персонализация,
- мультимодальность и интеграция.

1.6.5.1 Естественность и выразительность в синтезе речи

Искусственный интеллект в синтезе речи сегодня достиг высокого уровня естественности и выразительности. С помощью машинного обучения и алгоритмов обработки естественного языка, интеллектуальные системы могут производить высококачественную речь с натуральным интонационным окрасом и выражением.

Одним из главных достижений в этой области было создание глубоких нейронных сетей, которые обучаются на больших объёмах аудио– и текстовых данных, чтобы создавать высокоэффективные модели синтеза речи.

Научно–исследовательская и публикационная активность, связанная с направлением «Естественность и выразительность в синтезе речи» за последние 10 лет показана на рисунке 1.6.4 [1, 2].



Рисунок 1.6.4 – Научно–исследовательская и публикационная активность, связанная с направлением «Естественность и выразительность в синтезе речи» за последние 10 лет

В работе [71] продемонстрирован метод многоуровневого контроля просодии для синтеза выразительной речи. Новая модель синтеза выразительной речи, способна воспроизводить речь в определенном стиле, а также позволяет локально корректировать просодию сгенерированной речи. Модель основана на неавторегрессионной архитектуре «FastSpeech» и включает эталонный кодировщик для изучения глобальных вложений просодии и векторный квантованный вариационный автоэнкодер для создания мелкозернистых вложений просодии.

Актуальным направлением является преобразование текста в выразительный текст. Например, в работе [72] разработан метод преобразования контекстуального выразительного текста в речь. В работе было предложено использовать контекст для руководства процессом синтеза речи вместо того, чтобы полагаться на явные обозначения стилей и эмоций. Для решения этой задачи авторы создали синтетический набор данных и разработали эффективную структуру. Эксперименты показывают, что платформа может генерировать высококачественную выразительную речь на основе заданного контекста как в синтетических наборах данных, так и в реальных сценариях.

В работе [73] приведено исследование просодических особенностей моделирования вторичных эмоций, необходимых для синтеза эмпатической речи. Обычно вторичных эмоций в выразительном языке много, и, следовательно, разработка больших баз данных для каждой из вторичных эмоций обходится дорого. В работе доказана концепция синтеза речи со вторичными эмоциями с использованием ручного извлечения функций и моделирования этих функций с использованием низкоресурсоемкого подхода машинного обучения. С

использованием разработанных моделей получена эмоциональная система преобразования текста в речь для синтеза пяти вторичных эмоций: тревоги, извинения, уверенности, энтузиазма и беспокойства. Точность метода достигает 65 %.

1.6.5.2 Пользовательский голос и персонализация

Синтез речи с помощью искусственного интеллекта (ИИ) позволяет создавать голос, который может звучать как живой человек. Эта технология берет текст и преобразует его в звучащую речь, чтобы помочь людям, которые не могут говорить или имеют трудности с распознаванием голоса. Однако, чтобы голос был наиболее эффективным и личным, необходимо персонализировать его. Искусственный интеллект может учитывать различные параметры, такие как возраст, пол, географическое положение и особенности голоса пользователя, чтобы создать наиболее приятный и узнаваемый голос.

Искусственный интеллект в синтезе речи и персонализации голоса может иметь огромное значение для людей с ограниченными возможностями, а также для тех, кто хочет создать свой собственный уникальный голос для цифровых помощников или работать с рекламными роликами или аудиокнигами.

Научно–исследовательская и публикационная активность, связанная с направлением «Пользовательский голос и персонализация» за последние 10 лет показана на рисунке 1.6.5 [1, 2].



Рисунок 1.6.5 – Научно–исследовательская и публикационная активность, связанная с направлением «Пользовательский голос и персонализация» за последние 10 лет

В работе [74] продемонстрировано эмпирическое исследование с использованием лингвистических знаний о заполненных паузах для синтеза персонализированной спонтанной

речи. Для сравнительной оценки методов персонализированного ввода заполненной паузы и неперсонализированного прогнозирования заполненной паузы был разработан метод синтеза речи с неперсонализированным внешним предиктором заполненной паузы, обученным с помощью корпуса с несколькими динамиками. Данный метод вносит ясность в запутанность заполненных пауз в зависимости от позиции слова, то есть. необходимость точного прогнозирования позиций для естественности и необходимость точного прогнозирования слов для индивидуальности при оценке синтезированной речи.

В работе [75] разработан персонализированный синтез речи с остаточным управлением на основе изображения лица. В этой модели путем разработки двух речевых априорных значений вводится стратегия с остаточным управлением, чтобы направлять функцию лица, чтобы приблизиться к истинной речевой функции в обучении. Кроме того, учитывая ошибку абсолютных значений признаков и их направленное смещение, в работе формулируют новую функцию потерь с тремя элементами для кодировщика лиц. Экспериментальные результаты показывают, что речь, синтезированная нашей моделью, сравнима с персонализированной речью, синтезированной путем обучения большого количества аудиоданных в предыдущих работах.

В работе [76] продемонстрирован спонтанный синтез речи с тренировкой языково-речевой связности с использованием псевдозаполненных пауз. В работе был разработан тренинг лингвистической согласованности речи, который гарантирует согласованность языковых частей синтетической речи с заполненными паузами и без них. В предлагаемой тренировке последовательности используются не только паузы, заполненные реальной правдой, но и псевдопаузы. Эксперименты показывают, что этот метод улучшает естественность синтетической лингвистической речи и всей синтетической речи с предсказанием, заполнением, включением паузы.

1.6.5.3 Мультиmodalность и интеграция

Интеграция с другими технологиями и сервисами для синтеза речи в настоящее время имеет большую актуальность. Это связано с увеличением использования голосовых интерфейсов в разных сферах, таких как: медицина, автомобильная промышленность, домашняя автоматизация и др.

Интеграция с другими технологиями и сервисами позволяет расширить функциональность синтеза речи, улучшить его качество и снизить затраты на разработку. Например, с помощью интеграции с технологиями распознавания речи можно позволить пользователям взаимодействовать с устройством голосом. Интеграция со словарями определенных сфер знаний позволяет расширить словарь и улучшить качество синтеза речи в специализированных областях. Например, использование нейронных сетей может позволить

синтезировать более естественную речь и автоматически адаптировать ее к разным диалектам и акцентам.

Научно–исследовательская и публикационная активность, связанная с направлением «Мультимодальность и интеграция» за последние 10 лет показана на рисунке 1.6.6 [1, 2].



Рисунок 1.6.6 – Научно–исследовательская и публикационная активность, связанная с направлением «Мультимодальность и интеграция» за последние 10 лет

В работе [77] представлена мультимодальная сеть, основанная на аудиовизуальном внимании, для обнаружения фальшивых видео с говорящими лицами. В работе предлагается механизм аудиовизуального внимания для обнаружения более информативных функций, которые могут быть легко интегрированы в любую аудиовизуальную архитектуру CNN путем модульности. С дополнительным аудиовизуальным вниманием предлагаемая сеть «FTFDNet» способна повысить эффективность обнаружения в установленном наборе данных. Оценка работы показала отличную производительность при обнаружении поддельных видео говорящих лиц, которая может достигать уровня обнаружения 97 %.

В работе [78] разработан метод генерации говорящих голов, управляемый речевыми мимическими и звуковыми единицами на основе слияния мультимодальных представлений. В этом исследовании предлагается новая генеративная структура, которая содержит расширенную некаузальную временную сверточную сеть самовнимания в качестве модуля мультимодального слияния, чтобы способствовать изучению отношений кросс–модальных функций. Результаты количественных и качественных экспериментов показывают, что метод превосходит существующие методы как с точки зрения качества изображения, так и точности синхронизации губ.

В работе [79] продемонстрирован совместный анализ аудиовизуального синтеза речи с сенсорными измерениями для регулирования взаимодействия человека и робота. Аудиовизуальный синтез речи является одним из лучших способом общения между

человеком и роботом. Робот может общаться со своими пользователями, благодаря аудиовизуальной технологии синтеза речи. Слуховое восприятие зависит от использования массива микрофонов для обнаружения источников звука. Для этих объектов предлагается архитектура интеграции на основе коммуникационных протоколов с нисходящей иерархией. Верхний уровень интеграции содержит сообщение о количестве людей и связанных с ними состояний, которые отличаются от числа перцептивной сущности нижнего уровня.

1.6.6 Вывод

Основными тенденциями развития искусственного интеллекта в распознавании и синтезе речи являются глубокое обучение и нейронные сети, многоязычная и многодиалектная поддержка, интеграция с другими технологиями и сервисами в распознавании речи, пользовательский голос и персонализация в синтезе речи.

1.7 Искусственный интеллект в рекомендательных системах и интеллектуальных системах поддержки принятия решений

Рекомендательная система (далее РС) – это информационная система, которая автоматически находит и предлагает наиболее подходящие элементы для пользователя на основе его предыдущих действий или интересов.

Система поддержки принятия решений (далее СППР) – это когнитивная информационная система, целью которой является помощь лицам, принимающим решение в сложных условиях, путем как можно более полного и объективного анализа имеющихся данных для достижения целей принятия решения с учётом критериев эффективности и возможных ограничений, включая временные, на основе заложенных в систему алгоритмов в интерактивном режиме.

1.7.1 Искусственный интеллект в рекомендательных системах

Искусственный интеллект в рекомендательных системах – это использование алгоритмов машинного обучения и статистических методов для предоставления персонализированных рекомендаций пользователю на основе его предпочтений, поведения и истории взаимодействия с системой. Он использует данные о пользователе, товарах и контексте, чтобы предложить наилучшие варианты продуктов и услуг, учитывая индивидуальные потребности и предпочтения пользователя.

Основными тенденциями развития искусственного интеллекта в рекомендательных системах являются:

- объяснимость рекомендательных систем,
- гибридные рекомендательные системы,

— интеграция с Интернетом вещей (далее IoT., от англ. internet of things).

1.7.1.1 Объяснимость рекомендательных систем

Объяснимость рекомендательных систем – это способность системы предоставлять пользователю обоснование рекомендации, т.е. почему она рекомендует определенный продукт или услугу. Это понимание может быть в виде объяснения, основанного на данных, которые используются системой, или на основе алгоритмов, применяемых для генерации рекомендаций. Объяснимость рекомендаций помогает улучшить доверие пользователя к системе и может помочь ему сделать более осознанный выбор.

Научно–исследовательская и публикационная активность, связанная с направлением «Объяснимость рекомендательных систем» за последние 10 лет показана на рисунке 1.7.1 [1, 2].

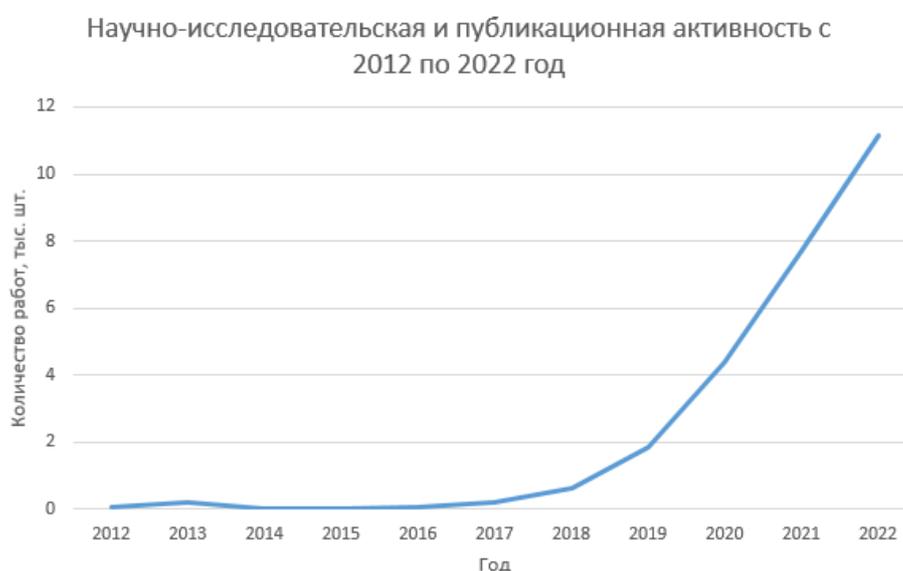


Рисунок 1.7.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Объяснимость рекомендательных систем» за последние 10 лет

В работе [80] разработана масштабируемая и объяснимая визуальная рекомендательная система SEV–RS. В работе сравнили «SEV–RS» с самыми современными моделями, такими как модель глубокого обучения с использованием сети «Graph Attention Network», на двух реальных наборах данных и одном синтетическом наборе данных. Результаты показывают, что «SEV–RS» дает более точные и объяснимые рекомендации. Кроме того, «SEV–RS» требует значительно меньшего вычислительного времени, чем сравниваемые модели глубокого обучения.

В работе [81] продемонстрировано выразительное моделирование скрытых функций для рекомендательных систем на основе объяснимой матричной факторизации. Разработанная структура системы объяснимых рекомендаций на основе функций использует только данные пользователя для объяснения рекомендации, не полагаясь на частные или внешние данные.

Рекомендации объясняются пользователю с помощью текстового сообщения и гистограммы. Предложенная модель была оценена с точки зрения точности предсказания рейтинга и разумности объяснения с использованием шести реальных наборов эталонных данных фильмов, книг, видеоигр и систем рекомендаций по моде. Результаты показывают, что предложенная модель может давать точные объяснимые рекомендации.

В работе [82] разработана система рекомендаций с графами знаний и с определением новизны, популярности и разнообразия объяснений. В этой работе было концептуализировано три новых свойства, которые моделируют качество объяснений (новизна взаимодействия, популярность общих объектов и разнообразие типов объяснений), и предложены подходы к ранжированию, способные оптимизировать эти свойства. Эксперименты с двумя общедоступными наборами данных показали, что разработанные подходы могут повысить качество объяснений в соответствии с предлагаемыми свойствами, справедливо для разных демографических групп, сохраняя полезность рекомендаций.

1.7.1.2 Гибридные рекомендательные системы

Гибридная рекомендательная система – это тип рекомендательной системы, который по-разному сочетает две или более стратегии рекомендаций, чтобы извлечь выгоду из их дополнительных преимуществ. Гибридные рекомендательные системы используются для создания и предоставления пользователям предложений по элементам и другим объектам с использованием различных стратегий. Гибридные рекомендательные системы подразделяются на разные типы в зависимости от того, как комбинируются стратегии рекомендаций.

Научно-исследовательская и публикационная активность, связанная с направлением «Гибридные рекомендательные системы» за последние 10 лет показана на рисунке 1.7.2 [1, 2].



Рисунок 1.7.2 – Научно-исследовательская и публикационная активность, связанная с

направлением «Гибридные рекомендательные системы» за последние 10 лет

В работе [83] разработана гибридная рекомендательная система «Eshop recommender» для интернет-магазина с использованием нечеткой экспертной системы. Рекомендательная система работает с несколькими параметрами: уровнем сходства с уже оцененными товарами, коэффициентом приобретенного товара и средним рейтингом товара. Разработанная система достигает результатов выше 90 % (точность, отзыв, F1-мера).

В работе [84] разработана гибридная рекомендательная система на основе глубокой нейронной сети с сетями пользователь-пользователь. Сети "Пользователь-пользователь" используются для обеспечения лучшей совместной работы и синергии в подобных моделях. В разработанном подходе объединение сетей "пользователь-пользователь" с глубокими нейронными сетями обеспечивает более высокую точность прогнозирования и лучшее время выполнения, чем другие современные методы. Обширные эксперименты с общедоступными наборами данных «Flixster» и «MovieLens» показывают, что разработанная методика превосходит текущие методы «Premier», достигая улучшения на 19 % в RMSE, на 9,2 % в MAE и на 4,1 % в результатах F1.

В работе [85] продемонстрирована гибридная рекомендательная система, основанная на поведении клиента и данных о транзакциях с использованием обобщенного алгоритма последовательного шаблона. На основе компонентов данных Алгоритм общения последней выборки создает правила в виде шаблона последовательности, которые затем используются для построения предложений продукта. Согласно результатам тестирования, предложения по продуктам, полученные на основе сочетания данных о транзакциях продаж и данных о поведении потребителей, превосходят рекомендации по продуктам, полученные только на основе данных о транзакциях продаж. Значения показателей точности, отзыва и «F-measure» выросли на 185,46, 170,83 и 178,43 %, соответственно.

1.7.1.3 Интеграция с Интернетом вещей

Рекомендательные системы в интернете вещей – это технологии, которые позволяют предоставлять персонализированные рекомендации для пользователей на основе их предпочтений и поведения при использовании устройств IoT. Это может включать в себя такие вещи, как умный дом, персональные устройства, умные гаджеты и другие подключенные устройства. Эти системы могут помочь улучшить опыт использования IoT, упростить принятие решений пользователями и повысить эффективность использования устройств.

Научно-исследовательская и публикационная активность, связанная с направлением «Интеграция с Интернет вещей» за последние 10 лет показана на рисунке 1.7.3 [1, 2].



Рисунок 1.7.3 – Научно–исследовательская и публикационная активность, связанная с направлением «Интеграция с Интернет вещей» за последние 10 лет

В работе [86] продемонстрирована гибридная система рекомендаций услуг Интернета вещей с использованием социального интернета вещей. Эффективную рекомендацию по услугам IoT можно получить на основе «Social IoT» (далее SIoT) с использованием данных, сгенерированных различными устройствами IoT, принадлежащими друзьям и друзьям друзей. В этой статье предлагается гибридный метод, который сочетает в себе неявную совместную фильтрацию и онтологию, чтобы рекомендовать пользователям персонализированные услуги IoT. Онтология используется для моделирования SIoT, где включаются социальные отношения между объектами в процесс рекомендаций наряду с рейтингами, в то время как совместная фильтрация прогнозирует рейтинги и генерирует рекомендации. Данный метод продемонстрировал свою эффективность и точность по сравнению с существующими алгоритмами.

В работе [87] разработана система рекомендаций на основе рейтингов, основанная на текстовых обзорах с использованием интеллектуальных устройств Интернета вещей. Данная работа посвящена повышению уровня точности текстовых обзоров с помощью системы рекомендаций при взаимодействии многочисленных пользователей с их доменами. Авторы внедрили систему рекомендаций с использованием пользовательского метода CF и представили важность совместной фильтрации в области фильмов с помощью системы рекомендаций. Весь эксперимент был проведён с использованием инструмента RapidMiner на основе Java. Точность данного метода сопоставима с аналогами, но эффективность гораздо выше.

В работе [88] рассмотрена оптимизация данных для промышленных систем рекомендаций на основе Интернета вещей. Было определено, что время выполнения может быть сокращено в среднем на 75 % при использовании десяти узлов, обменивающихся

оптимальным параметром декомпозиции, по сравнению с использованием одного. Было предложено использовать технологию «Spark» для ускорения вычислений в среднем на 20 % по сравнению с «Hadoop». Была предложена архитектура системы IoT, которая использует модифицированный алгоритм «Funk SVD» для оптимизации данных на пограничных устройствах и отслеживает эффективность предоставления рекомендаций с использованием центров управления и облачных ресурсов.

1.7.1.4 Доверенный искусственный интеллект

Надежный искусственный интеллект (далее НИИ) относится к разработке, развертыванию и использованию систем ИИ таким образом, чтобы гарантировать соответствие соответствующим законам, надежность и соблюдение этических принципов. НИИ характеризуется его способностью вызывать доверие у своих пользователей, таких как потребители, организации и общество. Он направлен на решение этических, юридических, социальных и технологических проблем, связанных с ИИ, и максимизацию его преимуществ при одновременном снижении рисков и опасностей.

Существует несколько измерений или принципов, которые обычно ассоциируются с заслуживающим доверия ИИ [89]:

— Прозрачность и объяснимость: системы искусственного интеллекта должны быть прозрачными в своих процессах принятия решений и предоставлять объяснения своих результатов. Это помогает пользователям понять, как система приходит к своим выводам, и укрепляет доверие к ее надежности и справедливости;

— Честный и беспристрастный: системы искусственного интеллекта должны быть спроектированы и обучены таким образом, чтобы избегать предвзятости и дискриминации. Они должны относиться ко всем отдельным лицам и группам справедливо и без фаворитизма, обеспечивая равные возможности и результаты;

— Надежные системы искусственного интеллекта должны быть спроектированы таким образом, чтобы они надежно работали в различных условиях и справлялись с неожиданными входными данными или ситуациями. Они должны быть способны поддерживать свою производительность и точность даже при столкновении с вызовами или враждебными атаками;

— Уважение к частной жизни: системы искусственного интеллекта должны защищать частную жизнь и конфиденциальность данных отдельных лиц. Они должны придерживаться соответствующих правил конфиденциальности и обеспечивать безопасную и ответственную обработку личной информации пользователей;

— Надежность: системы искусственного интеллекта должны разрабатываться с учётом требований безопасности для предотвращения несанкционированного доступа,

манипуляций или неправильного использования. В них должны быть предусмотрены меры предосторожности для защиты от киберугроз и обеспечения целостности и конфиденциальности системы и ее данных;

— Ответственный и подотчетный: системы искусственного интеллекта и их разработчики должны брать на себя ответственность за влияние своей технологии. Они должны нести ответственность за решения, принимаемые системой искусственного интеллекта, и предоставлять механизмы для устранения любых негативных последствий.

Научно–исследовательская и публикационная активность, связанная с направлением «Доверенный искусственный интеллект» за последние 10 лет показана на рисунке 1.9.9 [1, 2].



Рисунок 1.9.9 – Научно–исследовательская и публикационная активность, связанная с направлением «Доверенный искусственный интеллект» за последние 10 лет

В работе [90] предложен новый метод матирования видео с использованием нейронной сети, который не требует ввода данных от пользователей. Для обеспечения временной стабильности альфа-матов используется сглаживание выходных данных алгоритма сегментации изображения с помощью оценки движения и модулей сверточной LSTM на пропускаемых соединениях U-Net. Также предлагается алгоритм ложного движения, который позволяет обучать нейронную сеть на основе фотографий с достоверными альфа-матами и фоновыми видео, используя передний план оптического потока обучающего клипа в функциях потерь. Данный метод позволяет обходить необходимость в большом наборе аннотированных видеоданных и работать с видео, имитирующим движение, как в реальной жизни.

В работе [91] обсуждаются проблемы минимизации выпуклых функций, которые могут быть записаны в виде суммы различных оракулов, таких как полный градиент, производная по координате и стохастический градиент. В случае сильной выпуклости эти функции могут иметь различную сложность итерации методов первого порядка и количество вызовов оракула, необходимых для достижения заданной точности. Чтобы уменьшить количество

вызовов дорогих оракулов, авторы предлагают общую алгоритмическую структуру для разделения сложностей оракулов для каждой функции. Эта идея применима для различных случаев, включая конечную сумму и доступность через стохастический градиентный оракул. В результате мы получаем ускоренный случайный координатный спуск и ускоренный метод пониженной дисперсии с разделением сложности оракула.

В работе [92] улучшено использование гладкости более высокого порядка при оптимизации без производных.

В работе [93] рассматривается интеллектуальный анализ шаблонов, особенно анализ наборов элементов. Введено понятие структуры замыкания, основанной на закрытых наборах элементов и их минимальных генераторах, которые представляют содержимое набора данных с точки зрения закрытых множеств и классов эквивалентности наборов элементов. Предложена формализация структуры замыкания и ключей доступа в терминах анализа формальных понятий. Также представлен алгоритм GDPM для перечисления ключей доступа и обнаружения структуры замыкания, который возвращает характеристику содержимого набора данных с точки зрения уровней сложности. Эксперименты показывают, как можно практически использовать алгоритм GDPM и структуру замыкания.

В работе [94] обсуждается интеллектуальный анализ шаблонов и анализ наборов элементов, который позволяет анализировать бинарные наборы данных при поиске интересных и значимых наборов элементов и соответствующих правил ассоциации. Введено краткое представление — структуру замыкания — на основе закрытых наборов элементов и их минимальных генераторов для захвата внутреннего содержимого набора данных. В статье приводятся теоретические свойства паролей, которые являются краткими представителями закрытых множеств. Авторы предлагают формализацию структуры замыкания и ключей доступа в терминах анализа формальных понятий, которые хорошо приспособлены для изучения таких элементов. Также представлен алгоритм GDPM для перечисления ключей доступа и обнаружения структуры замыкания. Эксперименты показывают, как можно практически использовать алгоритм GDPM и структуру замыкания.

В работе [95] был использован подход Parenclitic Networks для исследования метилирования ДНК при синдроме Дауна. Были построены паренклитические сети и получены различные сигнатуры, которые обнаружили различия между людьми с СД и здоровыми людьми, различия между молодыми и пожилыми здоровыми людьми, различия между людьми с СД и их братьями и сестрами того же возраста, а также различия между СД и взрослым населением (их матерями). Были выявлены биологические процессы, связанные с фенотипом СД. Результаты исследования могут быть применены к другим заболеваниям. Работа была выполнена с помощью программного обеспечения с открытым исходным кодом,

которое доступно для всех исследователей.

В работе [96] было исследовано снижение когнитивных способностей с возрастом как основной проявление старения. Однако, количественные биомаркеры этого процесса и их соответствие различным биологическим часам до сих пор остаются открытой проблемой. Для изучения когнитивных показателей были использованы три теста, а именно: дифференцирование оттенков (кампиметрия); оценка арифметической правильности и обнаружение перевернутых букв. На основе их результатов были построены когнитивные часы, используя машинное обучение, которые с высокой точностью могут предсказывать хронологический, эпигенетический и фенотипический возраст. Интересно, что эти часы позволяют предсказывать хронологическую возраст с ошибкой 8,62 года. Исследование также выявило корреляцию между возрастными ускорениями и когнитивными способностями, что подтверждает глубокую связь между ними и статусом старения человека.

В работе [97] было проведено исследование оценки неопределенности в задачах обработки естественного языка. Большинство исследований на эту тему были проведены в области классификации изображений, поэтому данное исследование помогает заполнить пробел в данной области. Были рассмотрены современные методы UE для моделей Transformer и их эффективность была проверена на задачах распознавания именованных объектов и классификации текста. В результате этого были предложены две вычислительно эффективные модификации, одна из которых даже превосходит вычислительные интенсивные методы. Полученные результаты могут быть полезны для активного обучения, обнаружения ошибочной классификации, обнаружения выхода за пределы распределения и других задач.

В работе [98] представлен новый подход к извлечению открытой информации (OpenIE), который может быть более успешным, чем современные методы. Вместо того, чтобы итеративно извлекать триплеты, предлагается новый однопроходный метод, основанный на алгоритмах обнаружения объектов компьютерного зрения. Этот метод использует независимую от порядка потерю и архитектуру только на основе кодировщика для маркировки последовательностей, обеспечивающую уникальные прогнозы. Предлагаемая модель быстрее и демонстрирует превосходную или аналогичную производительность на стандартных эталонных тестах по сравнению с современными моделями. Эта модель устанавливает новый современный уровень производительности на CaRB, оцененный как OIE2016, и в 3,35 раза быстрее при выводе, чем предыдущий уровень техники. Также были оценены многоязычные версии модели в нулевой настройке для двух языков и показано улучшение производительности на 15 % на многоязычном Re-OIE2016, достигая 75 % F1 для португальского и испанского языков.

В работе [99] разработан ускоренный первично-двухградиентный метод для гладких и

выпукло-вогнутых седловых задач с билинейной связью.

В работе [100] представлен первый график размера шага для метода Ньютона, обеспечивающий быструю глобальную и локальную сходимость. Доказана глобальная скорость ($O(k^{1/2})$), соответствующая современным глобальным скоростям кубически регуляризованного метода Ньютона Поляка и Нестерова (2006), Ньютона Мищенко (2021), Дойков и Нестерова (2021). Также доказана локальная квадратичная скорость, соответствующая самой известной локальной скорости методов второго порядка. Формула размера шага проста, явна и не требует решения какой-либо подзадачи. Доказательства сходимости верны при предположениях об аффинной инвариантности, тесно связанных с понятием самосогласования. Метод имеет конкурентоспособную производительность по сравнению с существующими базовыми моделями, которые имеют такие же гарантии быстрой глобальной конвергенции.

В работе [101] описывается альтернативный алгоритм для ускорения оптимизации при сильновыпуклых задачах, основанный на методе Шамболя и Пока. Метод отличается от ProxSkip использованием неточного прокс-оператора градиентного метода и модифицированного шага двойного обновления для линейной сходимости. Общие результаты предлагают некоторые новые возможности для оптимизации сильновыпукло-вогнутых задач с билинейной связью, применимых для федеративного обучения. Метод требует меньше локальных шагов и выполняет детерминированное количество локальных шагов, чем ProxSkip. Также получены теоретические улучшения для оптимизации в подключенных сетях.

В работе [102] рассматриваются проблемы, связанные с использованием активного обучения в машинном обучении, и предлагает решения для уменьшения вычислительных ресурсов, необходимых для обучения моделей. Авторы предлагают два метода для задач классификации текста и тегирования, которые позволяют значительно сократить продолжительность итераций активного обучения. Кроме того, авторы демонстрируют, что их алгоритм, использующий псевдомаркировку и дистиллированные модели, может преодолеть препятствия, связанные с различиями между моделями сбора и последователей, используемых в активном обучении. Наконец, авторы показывают, что их алгоритм может обучать более выразительные модели с более высокой производительностью, несмотря на использование меньшей и более быстрой модели сбора данных.

представлен первый график размера шага для метода Ньютона, обеспечивающий быструю глобальную и локальную сходимость. Доказана глобальная скорость ($O(k^{1/2})$), соответствующая современным глобальным скоростям кубически регуляризованного метода Ньютона Поляка и Нестерова (2006), Ньютона Мищенко (2021), Дойков и Нестерова (2021). Также доказана локальная квадратичная скорость, соответствующая самой известной

локальной скорости методов второго порядка. Формула размера шага проста, явна и не требует решения какой-либо подзадачи. Доказательства сходимости верны при предположениях об аффинной инвариантности, тесно связанных с понятием самосогласования. Метод имеет конкурентоспособную производительность по сравнению с существующими базовыми моделями, которые имеют такие же гарантии быстрой глобальной конвергенции.

В работе [103] представлена эффективная стратегия запросов для метода активного обучения в абстрактном суммировании текста (ATS), основанная на принципах разнообразия, что позволяет уменьшить трудозатраты и улучшить производительность модели. Также показано, что использование самообучения может дополнительно повысить производительность модели. В целом, данная работа важна для развития методов машинного обучения, особенно в области обработки естественного языка.

В работе [104] исследовали распределенные задачи оптимизации в сети агентов без централизации. Полагалось, что функции потерь агентов подобны из-за подобия статистических данных или других факторов. Для уменьшения количества сообщений и достижения точности решения было предложено предварительно обусловленный, ускоренный распределенный метод. Численные результаты показали значительную экономию средств связи по сравнению со существующими схемами, особенно при решении плохо обусловленных задач.

1.7.2 Искусственный интеллект в интеллектуальных системах поддержки принятия решений

Искусственный интеллект в интеллектуальных системах поддержки принятия решений – это совокупность методов, алгоритмов и технологий, которые позволяют моделировать, анализировать и прогнозировать сложные процессы и явления в реальном времени, а также автоматически принимать оптимальные решения на основе анализа больших объёмов данных.

Основными направлениями развития искусственного интеллекта в интеллектуальных системах поддержки принятия решений являются:

- навигация, в том числе планирование маршрутов;
- планирование и принятие решений в промышленности;
- прогнозная аналитика.

1.7.2.1 Навигация, в том числе планирование маршрутов

Интеллектуальные системы поддержки принятия решений в навигации – это системы, которые используются для поддержки принятия решений в области навигации с помощью анализа и обработки больших объёмов данных, получаемых от различных источников,

включая спутники, камеры, сенсоры и другие приборы. Такие системы могут обрабатывать данные о погоде, морских течениях, состоянии трасс, местоположении и движении других судов и транспортных средств, а также других факторах, которые могут повлиять на безопасность и эффективность навигации.

Научно–исследовательская и публикационная активность, связанная с направлением «Навигация, в том числе планирование маршрутов» за последние 10 лет показана на рисунке 1.7.4 [1, 2].



Рисунок 1.7.4 – Научно–исследовательская и публикационная активность, связанная с направлением «Навигация, в том числе планирование маршрутов» за последние 10 лет

В работе [105] разработан механизм принятия решений и координации действий по предотвращению столкновения нескольких судов в сценариях смешанной навигации. Результаты моделирования показывают, что ближайшие расстояния между судами во всех сценариях не меньше предварительно определенного безопасного расстояния, и суда соответствуют Международным правилам предотвращения столкновений в море (далее МППС). В смешанных сценариях механизм уведомления о намерениях может адаптироваться к последствиям нарушения МППСС одним судном, чтобы уменьшить сложность операции и частоту изменения курса.

В работе [106] разработана интеллектуальная система поддержки принятия решений (далее ИСППР) для грузовых автомобильных перевозок. Эта ИСППР состоит из трех основных модулей, которые можно использовать по отдельности или вместе для: определения географической кластеризации транспортных услуг; предложение водителя транспорта; а также оптимизация маршрута и загрузки грузовиков. ИСППР была полностью спроектирована и разработана для поддержки данных в режиме реального времени и состоит из комплексного решения (E2ES), учитывая, что оно охватывает все основные транспортные и логистические

процессы с момента регистрации в базе данных до оптимизированного транспортного плана. Весь набор функций, включенных в IDSS, был разработан и проверен экспертами сектора грузовых перевозок из различных компаний, которые будут использовать предлагаемую систему.

В работе [107] продемонстрирован интеллектуальный подход к планированию пути автономных мобильных роботов на основе адаптивной системы нейро–нечеткого вывода. Моделирующий мобильный робот в «V-REP» был интегрирован в контроллер «ANFIS», закодированный в MATLAB. Результаты моделирования показывают, что предложенный метод планирования пути на основе функции утилиты «ANFIS» превосходит некоторые из связанных алгоритмов с точки зрения поиска путей, близких к оптимальным.

1.7.2.2 Интеллектуальные системы планирования и принятия решений в промышленности

Интеллектуальные системы планирования и принятия решений в промышленности – это программные продукты для автоматизации ряда рутинных операций и сбора данных в изделиях промышленности, на основе которых система может выполнять анализ данных и предоставлять пользователю рекомендации для принятия решений. Они предоставляют множество возможностей для оптимизации работы производственных процессов и повышения эффективности использования ресурсов. Использование интеллектуальных систем позволяет сократить время на принятие решений, снизить риск ошибок и улучшить качество продукции.

Научно–исследовательская и публикационная активность, связанная с направлением «Интеллектуальные системы планирования и принятия решений в промышленности» за последние 10 лет показана на рисунке 1.7.5 [1, 2].



Рисунок 1.7.5 – Научно–исследовательская и публикационная активность, связанная с направлением «Интеллектуальные системы планирования и принятия решений в

промышленности» за последние 10 лет

В работе [108] представлена двойная система интеллектуального планирования процесса производства авиационных деталей, основанная на данных и знаниях. Предлагается общая структура для системы интеллектуального планирования авиационных деталей, основанной на двойных данных и знаниях, и анализируются четыре стандартные процедуры, которые поддерживают вышеупомянутую справочную систему, а именно цифровую двойную модель процесса объединения данных, базу знаний динамического процесса, процесса принятия решений и оценки, прогнозирования качества обработки и оптимизации обратной связи процесса. Сконструированный таким образом испытательный стенд демонстрирует осуществимость и эффективность предлагаемого подхода.

В работе [109] продемонстрирован многокритериальный интеллектуальный метод принятия решений для многоступенчатого размещения блока управления электропитания на предприятиях электросетевого хозяйства. В работе предложен улучшенный алгоритм с двумя архивами для решения модели оптимизации размещения с пятью целями. Кроме того, предлагается нечеткий метод принятия решений, сочетающий субъективное и объективное, чтобы помочь предприятиям электросетевого хозяйства выбрать наиболее подходящее решение. Предложенный метод протестирован на нескольких шинных системах IEEE и польской системе 2383–bus, и результаты тестирования подтверждают его эффективность.

В работе [110] разработана цифровая интеллектуальная система обслуживания продуктов и тематическое исследование горнодобывающей промышленности. Подсистема интеллектуального обслуживания основана на данных и цифровых технологиях и обеспечивает диагностику неисправностей и онлайн–сервисное обслуживание сложных продуктов горнодобывающей промышленности. Исследование показывает, что все заинтересованные стороны могут беспрепятственно участвовать в процессе проектирования. Подсистема интеллектуальных продуктов использует итеративную оптимизацию (более 100 итераций) для интерактивного получения результатов проектирования. Подсистема интеллектуального обслуживания предоставляет цифровые услуги на протяжении всего процесса. Таким образом, обеспечивается стабильное, надежное и комплексное решение в области продуктов и услуг для сложных условий добычи полезных ископаемых. Выходные данные используются для руководства проектированием, отладкой и эксплуатацией физического оборудования. MSPSS отличается более высоким качеством и эффективностью проектирования, более коротким временем проектирования и более низкой стоимостью проектирования, чем традиционный метод проектирования.

1.7.2.3 Прогнозная аналитика

Интеллектуальные системы планирования и принятия решений используются в

прогнозной аналитике для создания более точных прогнозов и оптимизации процесса принятия решений. Эти системы используют различные алгоритмы и методы, включая интеллектуальный анализ данных, машинное обучение и искусственный интеллект. Одна из основных задач интеллектуальных систем планирования и принятия решений в прогнозной аналитике – предсказание будущих тенденций и трендов. Для этого системы используют анализ исторических данных, построение статистических моделей и оценку возможных сценариев развития событий.

Научно–исследовательская и публикационная активность, связанная с направлением «Прогнозная аналитика» за последние 10 лет показана на рисунке 1.7.6 [1,2].



Рисунок 1.7.6 – Научно–исследовательская и публикационная активность, связанная с направлением «Прогнозная аналитика» за последние 10 лет

В работе [111] продемонстрирована прогнозная аналитика для прогнозирования инсульта с использованием машинного обучения и нейронных сетей. В работе пришли к выводу, что возраст, заболевания сердца, средний уровень глюкозы и артериальная гипертензия являются наиболее важными факторами для выявления инсульта у пациентов. Кроме того, нейронная сеть персептрона, использующая эти четыре атрибута, обеспечивает наивысшую степень точности и наименьшую частоту промахов по сравнению с использованием всех доступных входных функций и других алгоритмов сравнительного анализа. Разработанная интеллектуальная система поддержки решения имеет точность, достигающую 78 %.

В работе [112] разработан метод аналитики социальных сетей с использованием больших данных для анализа конкуренции в системе принятия бизнес–решений. В настоящее время компании могут оценивать своих конкурентов в режиме реального времени и корректировать цены в соответствии с динамикой на рынке, анализировать неблагоприятные отзывы конкурентов, анализировать их и принимать решения. Предлагаемый метод исследует

влияние анализа социальных сетей на различные области, такие как недвижимость, организации и финансы. Разработанная модель помогает принимать более эффективные маркетинговые решения и разрабатывать уникальные стратегические подходы.

В работе [113] также разработан метод анализ данных социальных сетей для системы принятия бизнес-решений (далее СПБР) и конкурентного анализа. СПБР концентрируется на маркетинге и описывает операционный подход к получению ценной информации из социальных данных. СПБР выполняет краткое и точное описание текущих сценариев использования на основе фактических данных в соответствии с рекомендациями по принятию решений и инвестиционными возможностями, которые компании получают при использовании анализа социальных данных. Результаты экспериментов показывают, что СПБР достигает самых высоких показателей среди аналогов: точность 93,7 %, надежность 86,8 %, измерение F-17 %, частота отклонений 85,5 %.

1.7.2.4 Квантовые вычисления

Квантовые вычисления могут помочь в ускорении работы ИИ и повышении производительности, улучшении обработки данных, а также в создании более точных и надежных прогнозов в системах поддержки принятия решений.

Научно-исследовательская и публикационная активность, связанная с направлением «Прогнозная аналитика» за последние 10 лет показана на рисунке 1.7.7 [1, 2].



Рисунок 1.7.7 – Научно-исследовательская и публикационная активность, связанная с направлением «Прогнозная аналитика» за последние 10 лет

В работе [114] рассматривается проблема обработки больших объёмов данных в коммерческих организациях, которая станет невозможной к 2027 году при использовании традиционных компьютерных систем и алгоритмов искусственного интеллекта. Рассматривается возможное решение проблемы с помощью квантового вычисления и

предлагается использование алгоритма ECQ-KNN для обработки больших объёмов данных. В статье также представ алгоритм оптимизации CSO, который помогает улучшить производительность ECQ-KNN. Авторы провели сравнительный анализ производительности ECQ-KNN и традиционных методик и пришли к выводу о высокой эффективности предложенной системы.

В работе [115] посвящена решению проблемы энергосистем, а именно возрастающей вычислительной сложности, которую можно решать с помощью квантовых вычислений. Квантовые вычисления позволяют более быстро и эффективно решать задачи оптимизации, моделирования и машинного обучения. В статье представлен обзор существующих исследований и прогресса в разработке аппаратного обеспечения, программного обеспечения и алгоритмов для применения квантовых вычислений в различных областях, включая умные сети.

В работе [116] продемонстрирована интеграция квантовых вычислений в системы поддержки принятия решений. Интеграция квантовых вычислений в системы поддержки принятия решений может расширить их возможности и обеспечить более точное и эффективное принятие решений. Квантовые вычисления могут обрабатывать большие наборы данных и выполнять сложные вычисления, обеспечивая более точные прогнозы и решения. Используя вычислительные преимущества квантовых вычислений, системы поддержки принятия решений могут преодолеть ограничения классических вычислений и добиться значительного прогресса в процессах принятия решений

1.7.3 Вывод

Основными тенденциями развития искусственного интеллекта в рекомендательных системах и интеллектуальных системах поддержки принятия решений являются объяснимость рекомендательных систем, гибридные рекомендательные системы, интеграция с Интернет вещей, планирование и принятие решений в промышленности, прогнозная аналитика.

1.8 Генеративный искусственный интеллект

Генеративный искусственный интеллект – это тип искусственного интеллекта, который включает в себя системы и алгоритмы, которые могут создавать или генерировать новый и оригинальный контент или идеи, а не просто реагировать на ввод или следовать заранее запрограммированным правилам.

1.8.1 Генеративно–сопоставительные сети

Генеративно–сопоставительные сети (далее GAN, от англ. Generative adversarial networks)

представляют собой тип модели глубокого обучения, состоящей из двух нейронных сетей: генератора и дискриминатора. Генератор обучен генерировать синтетические данные, которые напоминают реальные данные, а дискриминатор обучен различать реальные и синтетические данные.

Две сети обучаются вместе, где генератор пытается создать данные, которые могут обмануть дискриминатор, а дискриминатор пытается правильно определить, являются ли данные реальными или синтетическими. В процессе состязательного обучения генератор учится создавать все более реалистичные данные, в то время как дискриминатор лучше различает реальные и синтетические данные.

Основными тенденциями развития генеративно–состязательных сетей являются:

- Генерация контента в компьютерной графике и дизайне,
- Генерация отчетов и документации в медицине, а также медицинская диагностика,
- Генеративно–состязательных сети ценообразовании и маркетинге.

1.8.1.1 Генерация контента в компьютерной графике и дизайне

GAN применяются в различных областях компьютерной графики и дизайна. Например, в архитектуре GAN использовались для создания макета дома и следования определенным установленным правилам [117, 118]. В дизайне продукта GAN применялись для синтеза дизайна [119]. Кроме того, GAN использовались для генерации изображений, передачи стилей и увеличения данных в графике и обработке изображений [120, 121].

Научно–исследовательская и публикационная активность, связанная с направлением «Генерация контента в компьютерной графике и дизайне» за последние 10 лет показана на рисунке 1.8.1 [1, 2].



Рисунок 1.8.1 – Научно–исследовательская и публикационная активность, связанная с

направлением «Генерация контента в компьютерной графике и дизайне» за последние 10 лет

В работе [122] разработан адаптивный алгоритм шумоподавления обучающего изображения, основанный на извлечении собственных значений и модели GAN. Алгоритм использует особенности изображения для предварительной обработки изображения и извлечения эффективной информации из изображения. Затем пограничный сигнал классифицируется в соответствии с пороговым значением, чтобы устранить проблему "чрезмерного подавления", а затем пограничный сигнал изображения извлекается для усиления эффективного сигнала в высокочастотном сигнале. Алгоритм использует адаптивную обучающую модель GAN для дальнейшего обучения изображения. Каждая итерация сети генераторов состоит из трех этапов. Как результат, получается наилучшее значение. Алгоритм статьи сравнивается с традиционным алгоритмом и литературным алгоритмом. При тех же условиях алгоритм может обеспечить эффективность работы при более высокой точности, и в то же время он все еще может шумоподавлять. Краевой сигнал изображения сохраняется и имеет лучший визуальный эффект.

В работе [123] продемонстрирован новый подход к созданию визуальных стимулов на основе искусственного интеллекта для разработки концепции окружающей среды. Путем тестирования шести вариантов классической модели GAN, обученных с помощью самодельного набора данных, были выбраны четыре подходящие модели для создания черно-белых миниатюр в качестве стимулов с захватывающей атмосферой. Кроме того, было проведено качественное исследование результатов четырех моделей с привлечением восьми экспертов в дизайне. Как было выяснено, данный подход действительно может упростить множество рабочих этапов дизайнеров.

В работе [124] продемонстрирован реалистичный синтез текста в изображение с генерирующими внимание состязательными сетями. Разработанный метод прост в реализации и практичен: выбор наиболее подходящих векторов слов и использование этих векторов для создания связанных подобластей изображения. Прототип в его нынешнем состоянии генерирует изображения птиц для приложения дизайна. Принимая во внимание результаты юзабилити-тестирования, команда разработчиков в будущих итерациях приложения надеется улучшить разрешение генерируемого изображения. Они планируют предоставить выбор для создаваемого множества изображений с дальнейшими улучшениями алгоритма генерации изображений.

1.8.1.2 Генерация отчетов и документации в медицине, а также медицинская диагностика

Научно-исследовательская и публикационная активность, связанная с направлением «Генерация отчетов и документации в медицине, а также медицинская диагностика» за

последние 10 лет показана на рисунке 1.8.2 [1, 2].



Рисунок 1.8.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Генерация отчетов и документации в медицине, а также медицинская диагностика» за последние 10 лет

В работе [125] разработан метод для генерации синтетического сигнала сердечного ритма (далее СССР) плода в условиях ограниченных ресурсов на основе GAN. Предлагаемый метод всесторонне оценивается с использованием 200 реальных записей СССР с четырех аспектов: эффективность тренировки, точность и разнообразие сгенерированных данных и потенциальное улучшение модели классификации. По сравнению с обучением на наборах данных с небольшой выборкой и наборами данных с несбалансированностью категорий, обучение на дополненных наборах данных повышает точность примерно на 12 % и 8 % соответственно. Разработанная архитектура обеспечивает эталонное значение для практического решения проблем с дисбалансом данных СССР и недостаточным количеством выборок.

В работе [126] продемонстрирована многомасштабная состязательная сеть для улучшения медицинских изображений «MAGAN». Метод объединяет многомасштабную информацию при извлечении признаков, создавая пирамиду признаков и фильтруя нерелевантную активацию, чтобы выделить важные области на основе распределения внимания, что положительно влияет на визуализацию. Более того, «MAGAN» усиливает ограничения на качество улучшенного изображения с точки зрения распределения освещения, деталей текстуры, глубоких семантических особенностей и плавности, чтобы повысить эффект улучшения. Экспериментальные результаты показывают, что по сравнению с шестью современными методами «MAGAN» обладает наиболее значительным эффектом улучшения изображения, а также лучше всего справляется с последующей задачей сегментации изображения.

В работе [127] проведено исследование по борьбе с COVID–19 с использованием

генеративных состязательных сетей и искусственного интеллекта для медицинских изображений. Исследования показали, что GAN обладают большим потенциалом для решения проблемы нехватки данных для изображений легких при COVID–19. Данные, синтезированные с помощью GAN, были полезны для улучшения обучения моделей сверточной нейронной сети, подготовленных для диагностики COVID–19. Кроме того, GAN также внесли свой вклад в повышение производительности сверточной нейронной сети за счет сверхразрешения изображений и сегментации.

1.8.1.3 Генеративно–состязательных сети ценообразовании и маркетинге

GAN – это тип алгоритма машинного обучения, который можно использовать в ценообразовании и маркетинге для создания синтетических данных для тестирования моделей и создания персонализированного контента. Сети GAN требуют больших вычислительных ресурсов, но успешно генерируют реалистичные изображения и данные временных рядов. Разработчики могут использовать популярные платформы глубокого обучения и предварительно обученные модели для реализации GAN в своих приложениях.

В ценообразовании GAN использовались для генерации синтетических данных, которые можно использовать для тестирования финансовых моделей и торговых стратегий [128]. GAN также использовались для моделирования цен на несколько классов активов для акций, фьючерсов и криптовалют [129].

Научно–исследовательская и публикационная активность, связанная с направлением «Генеративно–состязательных сети ценообразовании и маркетинге» за последние 10 лет показана на рисунке 1.8.3 [1, 2].

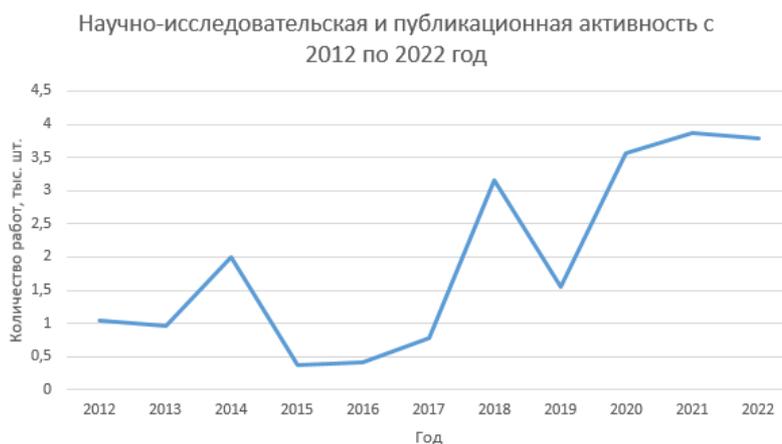


Рисунок 1.8.3 – Научно–исследовательская и публикационная активность, связанная с направлением «Генеративно–состязательных сети ценообразовании и маркетинге» за последние 10 лет

В работе [130] продемонстрирована технология прогнозирования цен на акции с помощью глубокой сверточной порождающей состязательной сети. Выполняя как

одноэтапное, так и многоэтапное прогнозирование, в работе наблюдают, что предлагаемая модель работает лучше, чем стандартные широко используемые инструменты, предполагая, что глубокое обучение (и, в частности, GAN) является многообещающей областью для прогнозирования финансовых временных рядов.

В работе [131] разработана технология генерация рекламных предложений с использованием усиленной генеративно–сопоставительной сети. Неспособность быстро поэкспериментировать, чтобы найти то, что работает, приводит к неудовлетворенности пользователей и трате маркетингового бюджета. Для быстрого эксперимента в работе решили автоматически генерировать рекламные предложения. Процесс создания рекламного объявления с заданной целевой страницы рассматривается как проблема обобщения текста, и был принят подход к обобщению абстрактного текста. Как показывают результаты, данный подход действительно может сократить как затраты как временный, так и денежные.

В работе [132] разработан алгоритм гибридного прогнозирования многомодульной генеративно–сопоставительной сети (далее MMGAN–HPA, от англ. Multi–Model Generative Adversarial Network Hybrid Prediction Algorithm) для прогнозирования цен на фондовом рынке. GAN мало используются для прогнозирования фондового рынка из–за сложности установки правильного набора гиперпараметров. В этой статье эта проблема преодолевается с помощью обучения с подкреплением и байесовской оптимизации. Фреймворк глубокого обучения, основанный на GAN, названный Stock–GAN, реализован с генератором и дискриминатором. Первый реализуется с помощью LSTM, варианта рекуррентной нейронной сети, а второй использует сверточную нейронную сеть. Предлагается алгоритм, названный алгоритмом гибридного прогнозирования на основе генеративно–сопоставительной сети (далее GAN–HPA, от англ. Hybrid Prediction Algorithm). Эмпирическое исследование показало, что Stock–GAN обеспечивает многообещающие результаты в прогнозировании цен на акции по сравнению с современной моделью, известной как гибридный алгоритм прогнозирования на основе нескольких моделей (далее MM–HPA, от англ. Multi–Model based Hybrid Prediction Algorithm). Впоследствии MM–HPA и GAN–HPA объединились, чтобы сформировать еще одну гибридную модель, известную как MMGAN–HPA, для повышения производительности по сравнению с MM–HPA и GAN–HPA.

1.8.2 Поточковые модели

Поточковые модели в генеративном искусственном интеллекте определяются как алгоритмические модели, которые используют поток данных для генерации новых объектов или решения задач. В этих моделях используется ряд математических функций, известных как преобразования на основе потока, для сопоставления входных данных с новым распределением выходных данных.

Основными тенденциями развития потоковых моделей являются:

- Обработка и генерация больших объёмов данных в здравоохранении,
- Оптимизация и ускорение процессов в промышленности,
- Оптимизация в розничной торговле.

1.8.2.1 Обработка и генерация больших объёмов данных в здравоохранении

Потоковые модели работают как нейронные сети, но вместо того, чтобы обрабатывать данные целиком, они разбивают их на маленькие кусочки и обрабатывают каждый из них по отдельности. Это позволяет потоковой модели обучаться на большом объёме данных и быстро возвращать результаты.

В здравоохранении потоковые модели могут использоваться для анализа больших объёмов медицинских данных, таких как изображения, результаты тестов и т.д. При этом модель может обрабатывать данные в реальном времени, что позволяет быстро и точно проводить диагностику и принимать решения о лечении.

Научно–исследовательская и публикационная активность, связанная с направлением «Обработка и генерация больших объёмов данных в здравоохранении» за последние 10 лет показана на рисунке 1.8.4 [1, 2].



Рисунок 1.8.4 – Научно–исследовательская и публикационная активность, связанная с направлением «Обработка и генерация больших объёмов данных в здравоохранении» за последние 10 лет

В работе [133] продемонстрирована генеративная потоковая модель для увеличения объёмных данных в трехмерном глубоком обучении для компьютерной томографической колонографии. В работе была разработана 3D–сверточную нейронную сеть (далее 3D CNN) на основе потоковой генеративной модели (далее 3D Glow) для генерации синтетических объёмов, представляющих интерес (далее VOI), которая имеет характеристики, аналогичные характеристикам VOI из его обучающего набора данных. 3D Glow был обучен генерировать синтетические VOI полипов с использованием нашей коллекции клинических случаев

компьютерной томографии колонографии. Оценка была выполнена с использованием исследования «human observer» с участием трех наблюдателей и исследования классификации полипов на основе «CAdE» с использованием «3D DenseNet».

В работе [134] рассмотрен гибкий подход AutoML, который позволяет автоматизированное машинное обучение для классификации временных рядов. Предложенный подход использует несколько стратегий генерации признаков и классификаторы для конвейера моделирования. Генерация пайплайна основана на эволюционном алгоритме, что позволяет получить более надежное и легкое решение задачи классификации. Сравнение с подходами с наиболее высокими результатами было проведено в архиве UEA/UCR, что подтверждает высокое качество предлагаемого подхода. Этот подход может быть использован как в производственном пайплайне обработки данных, так и как инструмент моделирования и изучения свойств природных процессов.

В работе [135] разработана комплексная мультимодальная генерация рентгеновских снимков. Обширные эксперименты с данными реальных рентгеновских отчетов показывают, как увеличение данных с использованием синтезированных мультимодальных образцов может повысить производительность различных контролируемых задач, включая рентгеновскую классификацию COVID-19 с ограниченным количеством образцов. Рентгенологи также подтверждают качество сгенерированных изображений и отчетов. В работе количественно показывают, что синтетические наборы данных, сгенерированные EMIXER, могут улучшить классификацию рентгеновских изображений и модели создания отчетов, чтобы достичь улучшения на 5,94 % и 6,9 % по сравнению с моделями, обученными только на реальных образцах данных.

В работе [136] рассмотрены синтетические данные в здравоохранении. Хотя объем данных в здравоохранении растет экспоненциально, создание наборов данных для новых задач и отражающих разнообразный набор условий и причинно-следственных связей не является тривиальным. Кроме того, эти данные очень чувствительны и часто специфичны для конкретного пациента. Недавние исследования начали иллюстрировать потенциал синтетических данных во многих областях медицины, но систематического обзора литературы не существует. В этой статье представляются примеры физического и статистического моделирования для создания данных и предлагаемые приложения в здравоохранении и медицине.

1.8.2.2 Оптимизация и ускорение процессов в промышленности

Генеративный искусственный интеллект и потоковые модели находят применение в различных областях промышленности, включая производство, энергетику, транспорт, здравоохранение и многие другие. Одной из основных проблем, с которой сталкиваются

производственные предприятия, является оптимизация производственных процессов, в том числе управление потоками материалов и заказов. Использование генеративного искусственного интеллекта и потоковых моделей позволяет значительно ускорить и оптимизировать процессы в производстве, увеличить качество продукции и снизить затраты на производство.

Научно–исследовательская и публикационная активность, связанная с направлением «Оптимизация и ускорение процессов в промышленности» за последние 10 лет показана на рисунке 1.8.5 [1, 2].



Рисунок 1.8.5 – Научно–исследовательская и публикационная активность, связанная с направлением «Оптимизация и ускорение процессов в промышленности» за последние 10 лет

В работе [137] разработана метагенеративная состязательная сеть на основе градиентного потока (далее МССНОГП) для увеличения данных при диагностике неисправностей. Хорошо обученный МССНОГП может создавать достаточное количество имитированных образцов. После этого качество сгенерированных данных оценивается тремя методами, а именно анализом спектра на основе быстрого преобразования Фурье, визуализацией признаков на основе t -распределенного стохастического встраивания соседей и новыми грамианскими угловыми полями суммирования. на основе формирования изображения сигнала. В конечном счете, эти вновь сгенерированные сигналы вибрации используются в качестве дополнительных данных для обучения модели классификации и завершения последующих задач диагностики. Обширные эксперименты иллюстрируют эффективность и превосходство предложенной структуры над другими современными методами.

В работе [138] продемонстрировано сквозное обнаружение неисправностей с использованием потоковой модели. Обнаружение неисправностей широко изучалось как в научных кругах, так и в промышленности. Редкость ошибочных выборок в реальном мире ограничивает использование многих контролируемых моделей, а зависимость от экспертных

знаний предметной области для разработки признаков создает другие барьеры. Предлагаемый подход был проверен на двух наборах данных сигналов вибрации; он оказался лучше нескольких альтернатив.

В работе [139] разработана модель прогнозирования сценария электроснабжения жилого дома на основе модели генерации переносимого потока. Сетевые параметры части ступени структуры потока были заморожены в предобучающей модели, структурные параметры размороженной ступени структуры потока были доработаны и обучены по данным об электропотреблении домохозяйств в целевой области. Экспериментальные результаты показывают, что алгоритм в сочетании с переносом модели хорошо справляется с задачей прогнозирования нагрузки на жилье для небольших выборок.

1.8.2.3 Оптимизация в розничной торговле

Потоковая передача данных стала важнейшим фактором развития розничной торговли, учитывая эволюцию многоканального взаимодействия с клиентами и персонализированные рекомендации. На данный момент актуальными применениями потоковых моделей в розничной торговле являются: оптимизация складских процессов, использования торгового оборудования (стеллажей, касс, торговая мебель и др.), анализ и управление запасами и наличия товара, а также прогнозирования спроса.

Научно–исследовательская и публикационная активность, связанная с направлением «Оптимизация в розничной торговле» за последние 10 лет показана на рисунке 1.8.6 [1, 2].



Рисунок 1.8.6 – Научно–исследовательская и публикационная активность, связанная с направлением «Оптимизация в розничной торговле» за последние 10 лет

В работе [140] продемонстрирована модель для оценки дизайна и эстетики продукта с помощью генеративного ИИ с использованием потоковых моделей. Разработанная модель хорошо предсказывает привлекательность новых эстетических дизайнов – улучшение на 43,9 % по сравнению с базовым уровнем и существенное улучшение как по сравнению с обычными моделями машинного обучения, так и с предварительно обученными моделями

глубокого обучения. Новые конструкции автомобилей генерируются контролируемым образом.

В работе [141] разработана модель ценообразования с использованием генеративных сетей на основе потоков. Разработанная модель может эффективно моделировать случайные состояния за короткое время. Кроме того, они могут предоставить явные функции плотности вероятности для моделей стохастической волатильности благодаря уникальному преимуществу генеративных моделей на основе потоков. Это приводит к довольно точным ценам опционов, достигаемым путем моделирования случайных состояний с помощью сети или интегрирования выплат опционов для сетевой плотности.

В работе [142] разработан метод распознавания окклюзированных товаров в соответствии с предварительным выводом на основе генеративно–состязательной сети. Экспериментальные результаты показывают, что пиковое отношение сигнал–шум и структурное сходство при улучшенном предварительном выводе на 0,7743 и 0,0183 выше, чем у других моделей соответственно. По сравнению с другими оптимальными моделями mAP повышает точность распознавания на 1,2 % и точность распознавания на 2,82 %. Это исследование решает две проблемы: одна – окклюзия, вызванная руками, а другая – высокое сходство товаров, что соответствует требованиям точности распознавания товаров в сфере интеллектуальной розничной торговли и демонстрирует хорошие перспективы применения.

1.8.3 Модели на базе трансформера

Модели на основе трансформеров – это нейронные сети, которые изучают контекст и значение, отслеживая отношения в последовательных данных, таких как текст, и применяют методы внимания или самоконтроля для обнаружения незаметных способов, которыми даже удаленные элементы данных в серии влияют друг на друга и зависят друг от друга [143].

Основными тенденциями развития модели на базе трансформера являются:

- Чат-боты в маркетинге и рекламе,
- Генерация текста и данных в здравоохранении,
- Анализ и прогнозирование в финансах и банковском деле.

1.8.3.1 Чат-боты в маркетинге и рекламе

Модели на базе трансформера обучаются на огромных объемах данных и используют передовые методы обработки естественного языка для понимания контекста, настроек и других важных факторов, влияющих на эффективность маркетинговых стратегий. Использование данной технологии приносит маркетологам лучший из доступных инструментов для понимания своих клиентов и создания персонализированные маркетинговые кампании, адаптированные к уникальным потребностям и предпочтениям

каждого человека. Это может привести к увеличению вовлеченности, повышению коэффициента конверсии и, в конечном итоге, к увеличению дохода.

Научно–исследовательская и публикационная активность, связанная с направлением «Чат-боты в маркетинге и рекламе» за последние 10 лет показана на рисунке 1.8.7 [1, 2].



Рисунок 1.8.7 – Научно–исследовательская и публикационная активность, связанная с направлением «Чат-боты в маркетинге и рекламе» за последние 10 лет

В работе [144] продемонстрирован метод масштабирования идентификации естественного языка (далее NLI, от англ. natural language inference) с помощью адаптеров трансформеров. Последнее время подход глубокой генерации, основанный на декодерах–трансформерах, превзошел своих аналогов и показал наилучшие результаты на тестовых наборах данных NLI. В работе исследуется этот подход, чтобы определить практические последствия по сравнению с традиционными современными системами NLI. Также были представлены адаптеры–преобразователи, чтобы устранить ограничения памяти и повысить скорость обучения/логического вывода для масштабирования приложений NLI. В перспективе данная технология способна устранить языковые барьеры в чат–ботах, а также в автоматизированной технической поддержке.

В работе [145] продемонстрировано исследование в области навигации по этической сфере технологии чат–ботов на основе GPT. В работе утверждается, что инструменты на основе ChatGPT могут помочь маркетологам создавать контент быстрее и потенциально с качеством, аналогичным создателям контента. Это также может помочь маркетологам проводить более эффективные исследования и лучше понимать клиентов, автоматизировать обслуживание клиентов и повышать эффективность. Также исследование затрагивает этические последствия и потенциальные риски для маркетологов, потребителей и других заинтересованных сторон, которые необходимы для маркетинга на основе ChatGPT.

В работе [146] продемонстрирована разработка классификации и обнаружения спам–сообщений с использованием встраивания на основе преобразователя и обучения ансамблю.

В предлагаемой модели используется метод встраивания текста, основанный на последних достижениях GPT–3. Этот метод обеспечивает высококачественное представление, которое может улучшить результаты обнаружения. Кроме того, был использован метод ансамблевого обучения, в котором четыре модели машинного обучения были сгруппированы в одну модель, которая работала значительно лучше, чем ее отдельные составные части. Экспериментальная оценка модели была выполнена с использованием набора данных для сбора SMS–спама. Полученные результаты показали современную производительность, которая превзошла все предыдущие работы с точностью до 99,91 %.

1.8.3.2 Генерация текста и данных в здравоохранении

Модели на основе трансформера продемонстрировали большой потенциал в приложениях для здравоохранения, особенно при обработке больших и сложных наборов данных и повышении производительности моделей. Также на данный момент активно разрабатываются модели на базе трансформера для приложений медицины, например, одной из многообещающих технологий является генерация текста и данных, а также медицинских отчетов. Однако остаются проблемы с получением обширных наборов данных, развертыванием моделей в клинических условиях и решением нормативных вопросов [147].

Научно–исследовательская и публикационная активность, связанная с направлением «Генерация текста и данных в здравоохранении» за последние 10 лет показана на рисунке 1.8.8 [1, 2].

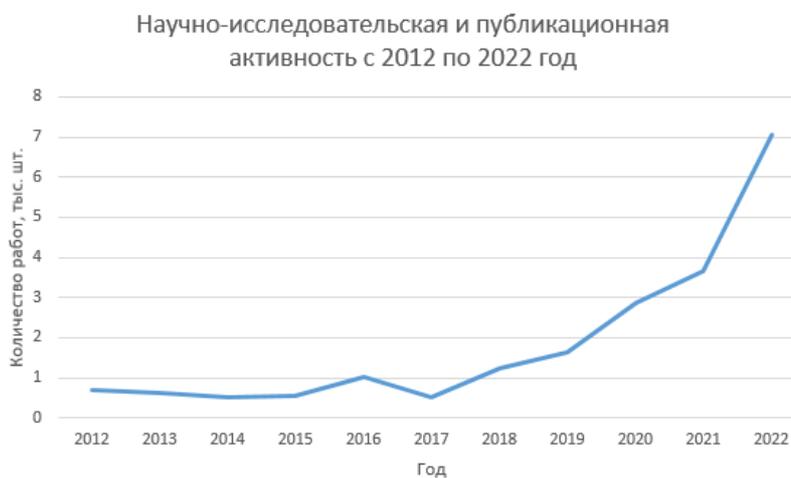


Рисунок 1.8.8 – Научно–исследовательская и публикационная активность, связанная с направлением «Генерация текста и данных в здравоохранении» за последние 10 лет

В работе [148] продемонстрирована оценка влияния полноты медицинских данных на глубинные генеративные модели. В работе были созданы как наборы данных поперечного сечения, так и панельные наборы данных с различной частотой пропусков и подмножеств и обучаем генеративные состязательные сети, вариационные автокодеры и авторегрессионные модели (трансформаторы) на основе этих наборов данных. Затем были сравнены

распределения сгенерированных данных с исходными данными обучения, чтобы измерить сходство. Было обнаружено, что повышенная неполнота напрямую коррелирует с увеличением различий между оригинальными и сгенерированными образцами.

В работе [149] рассмотрено внедрение предварительно обученных лингвистических моделей большого объёма с искусственным интеллектом в системе здравоохранения США: перспективы генеративного предварительно обученного GPT-3 в качестве модели обслуживания. По оценке аналитиков, вероятнее всего, одно из главных применений GPT-3 в ближайшие года в медицинской отрасли является ассистент и чат-бот для снижения нагрузки на клиники и обеспечения качественного обслуживания во время экстремальных и пиковых нагрузках. Запросы GPT-3 будут формулироваться путём сбора информации о пациенте и отправляются в API, где задачи распределяются и обрабатываются, производя выходные данные, которые затем отправляются в систему электронных медицинских записей и приложение на фронт-энд. Текст разговора также суммируется и сохраняется в медицинские записи пациентов.

В работе [150] рассмотрена диагностическая точность списков дифференциальной диагностики, сгенерированных чат-ботом GPT-3. Частота постановки правильного диагноза с помощью GPT-3 в пределах десяти дифференциально-диагностических списков составила 28/30 (93,3 %). Частота постановки правильного диагноза врачами по-прежнему была выше, чем у ChatGPT-3 в пределах пяти дифференциально-диагностических списков (98,3 % против 83,3 %, $p = 0,03$). Частота последовательных дифференциальных диагнозов среди врачей в десяти списках дифференциальных диагнозов, созданных ChatGPT-3, составила 62/88 (70,5 %). Таким образом, это исследование демонстрирует высокую диагностическую точность дифференциально-диагностических списков, созданных ChatGPT-3 для клинических случаев с общими основными жалобами. Это говорит о том, что чат-боты с искусственным интеллектом, такие как ChatGPT-3, могут генерировать хорошо дифференцированный список диагнозов для общих основных жалоб.

1.8.3.3 Анализ и прогнозирование в финансах и банковском деле

Модели на основе трансформаторов набирают популярность в финансах и банковском деле благодаря их способности анализировать большие наборы данных и давать точные прогнозы.

Научно-исследовательская и публикационная активность, связанная с направлением «Анализ и прогнозирование в финансах и банковском деле» за последние 10 лет показана на рисунке 1.8.9 [1, 2].



Рисунок 1.8.9 – Научно–исследовательская и публикационная активность, связанная с направлением «Анализ и прогнозирование в финансах и банковском деле» за последние 10 лет

В работе [151] показана применение технологии GPT в бухгалтерском учёте. Одной из перспективных направлений развития GPT в бухгалтерии является автоматизированная генерация бухгалтерской отчетности. Автора также затрагивают высокие риски, связанные с возможностью ошибок, и как следствие обвинений в подделывании бухгалтерской отчетности. Решением данной проблемы является дополнительная генеративная модель, которая будет проверять данные отчетности перед публикацией.

В работе [152] рассмотрены разные модели оценки анализа настроений. Были сравнены по производительности GPT с FinBERT, предварительно обученного BERT с финансовым блоком и доработанного с гораздо большим объёмом помеченных данных. Результаты показывают, что FinBERT достигает превосходной производительности в классификации настроений, несмотря на то, что имеет всего 110 миллионов параметров по сравнению со 175 миллиардами параметров GPT. В работе обнаружили, что и GPT, и FinBERT обеспечивают большую объяснительную силу, чем словари L&M, при прогнозировании доходности акций. Однако FinBERT превосходит GPT в прогнозировании реакции рынка на финансовые настроения. В целом исследование подчеркивает потенциал внедрения кратковременного обучения на основе LLMS в области бухгалтерского учёта и финансов, но, что более важно, объясняется масштабированность машинного обучения "больше – не обязательно лучше" и подчеркивается важность знаний, специфичных для конкретной предметной области, для задач, требующих более высокого уровня точности.

В работе [153] рассмотрено применение GPT в качестве финансового консультанта Davinci и ChatGPT набрали 58 % и 67 % баллов по тестированию на финансовую грамотность, соответственно, по сравнению с базовым показателем в 31 %. Однако люди переоценили

эффективность GPT (79,3 %), и в дилемме сбережений они в значительной степени полагались на советы GPT ($WOA = 0,65$). Снижение субъективных финансовых знаний привело к увеличению принятия рекомендаций. В работе также рассматривается риск чрезмерной зависимости от современных моделей большого языка и то, как может измениться их полезность для людей без профильного образования.

Также стоит выделить следующие технологии:

- Модели на основе трансформеров используются для создания уровней принятия решений на базе ИИ в банках [154];
- Модели на основе трансформеров используются для рекомендации товаров в банковской сфере [155];
- Модели на основе трансформеров используются для обнаружения мошенничества, рекомендаций по продуктам и семантического поиска в финтехе [156];
- Модели на основе трансформеров используются для прогнозирования дефолтов на корпоративных рынках [157];
- Модели на основе трансформеров могут быть использованы для оценки актуальности информации в социальных сетях [158];
- Модели на основе трансформера могут улучшить модели принятия кредитных решений в кредитных процессах, что приведет к увеличению доходов, снижению уровня кредитных убытков и повышению эффективности [159].

1.8.4 Вывод

Таким образом, основными тенденциями развития генеративного искусственного интеллекта являются генерация контента в компьютерной графике и дизайне, генерация отчетов и документации в медицине, медицинская диагностика с использованием генеративно-состязательных сетей, обработка и генерация больших объемов данных в здравоохранении, оптимизация и ускорение процессов в промышленности, оптимизация в розничной торговле с использованием потоковых моделей, чат-боты в маркетинге и рекламе, обработка и генерация больших объемов данных в здравоохранении, анализ и прогнозирование в финансах и банковском деле в моделях с использованием моделей на базе трансформера.

1.9 Автоматическое машинное обучение

Вопрос автоматического машинного обучения вынесен в отдельный подраздел, так как машинное обучение — это подмножество искусственного интеллекта, которое фокусируется на его обучении на данных без явного программирования. Это делается с помощью алгоритмов, которые могут автоматически улучшаться при наличии большего количества

данных. При этом искусственный интеллект — более широкое понятие, включающее все методы повышения интеллектуальности программного обеспечения, включая машинное обучение.

Автоматическое машинное обучение – это процесс автоматизации различных этапов машинного обучения, включая предобработку данных, выбор наилучших моделей и гиперпараметров, оптимизацию модели и ее оценку.

Автоматическое машинное обучение позволяет ускорить процесс разработки моделей машинного обучения и снизить требования к квалификации специалистов в этой области. Технология может быть реализована как в виде отдельных инструментов и библиотек, так и встроен в другие приложения и системы.

1.9.1 Автоматизированная генерация признаков

Автоматизированная генерация признаков – это процесс автоматического генерирования полезных признаков из исходных данных для использования в моделях машинного обучения. Автоматизированная разработка признаков может быть использована для улучшения точности и эффективности моделей машинного обучения и помогает сократить количество времени и усилий, необходимых для создания и развертывания этих моделей.

Основными тенденциями развития автоматизированной разработки признаков являются:

- глубокое обучение,
- генетические алгоритмы,
- автоэнкодеры.

1.9.1.1 Глубокое обучение

Глубокое обучение – это подмножество машинного обучения, в котором используются искусственные нейронные сети с репрезентативным обучением для имитации поведения человеческого мозга, что позволяет ему «учиться» на больших объёмах данных. Методы глубокого обучения используются для создания искусственных нейронных сетей с несколькими уровнями обработки для изучения представлений данных с несколькими уровнями абстракции.

Признаки глубоких нейронных сетях являются ключевым элементом, позволяющим модели извлекать высокоуровневые признаки из сырых данных. Для решения сложных задач, таких как обработка изображений, звуковых волн или естественного языка, требуется большое количество признаков, что может быть трудоемким и требовать большого количества экспертной работы.

Научно–исследовательская и публикационная активность, связанная с направлением «Глубокое обучение» за последние 10 лет показана на рисунке 1.9.1 [1, 2].



Рисунок 1.9.1 – Научно–исследовательская и публикационная активность, связанная с направлением «Глубокое обучение» за последние 10 лет

В работе [160] рассматривается использование методов машинного обучения для определения акустических параметров материалов в сложных случаях, также как и ускорение волнового конечно-элементного анализа. В частности, анализируется управляемый данными подход к обращению данных смещения к теоретической аппроксимации постоянной распространения и, следовательно, к параметрам материала с использованием символьной регрессии. Два примера применения этих методов рассматриваются на основе моделей волноводов: осевого стержня и кольцевой мембраны. Результаты показывают, что использование методов машинного обучения может быть эффективным инструментом для определения акустических параметров материалов в сложных случаях, а также ускорить процесс волнового конечно-элементного анализа.

В работе [161] продемонстрировано применение методов глубокого обучения в умном сельском хозяйстве. Точное сельское хозяйство использует передовые технологии, такие как IoT, интеллектуальный анализ данных, анализ данных, машинное обучение для сбора данных, обучения систем и прогнозирования результатов. В последнее время фермеры сталкиваются с различными проблемами, в том числе неурожаем из-за меньшего количества осадков, бесплодием почвы и так далее. В связи с изменениями, происходящими в окружающей среде, предлагаемая работа помогает определить, как разумно управлять посевами и собирать урожай. Результаты показывают, что методы глубокого обучения значительно повышают эффективность выращивания сельскохозяйственных культур.

В работе [162] продемонстрировано применение глубокого обучения в прецизионной и геномной медицине, что может улучшить медицинское обслуживание пациентов, позволяя системе рассуждать и учиться, улучшая процесс принятия решений врачом. Многие

характеристики клеток, в том числе активация генов, связывание белков с нуклеиновыми кислотами и сплайсинг, могут быть измерены с высокой пропускной способностью и использованы в качестве обучающих целей для прогностических моделей.

Также важной разработкой в данной области является работа [163], в которой разработана система обнаружения апноэ во сне по ЭКГ с помощью автоматического машинного обучения с глубоким обучением. Сначала сигналы электрокардиограммы были предварительно обработаны и сегментированы, а затем были применены методы машинного обучения и глубокого обучения для обнаружения апноэ во сне. Имеющиеся данные были разделены на обучающий набор для настройки параметров модели, проверочный набор для настройки гиперпараметров, предотвращения переобучения и улучшения обобщаемости моделей и тестовый набор для оценки обобщаемости моделей на невидимых данных. Затем эта процедура была повторена в пятикратной схеме перекрестной проверки, так что все записи были однажды обнаружены в тестовом наборе. Было обнаружено, что наилучшие характеристики обнаружения достигаются гибридными глубинными моделями, где наилучшие точность, чувствительность и специфичность составили 88,13 %, 84,26 % и 92,27 %, соответственно.

1.9.1.2 Генетические алгоритмы

Генетические алгоритмы – это метод оптимизации, идея которого заключается в использовании принципов естественного отбора и генетического кодирования для создания новых решений. Эти алгоритмы могут быть использованы для автоматизированной генерации признаков, проведения отбора наиболее важных признаков и оптимизации параметров моделей.

Научно–исследовательская и публикационная активность, связанная с направлением «Генетические алгоритмы» за последние 10 лет показана на рисунке 1.9.2 [1, 2].



Рисунок 1.9.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Генетические алгоритмы» за последние 10 лет

В работе [164] разработана автоматическая и интеллектуальная система визуализации контента на основе глубокого обучения и генетического алгоритма. Разработанная система использовалась для тестирования различного количества объектов обучения для различных областей науки. Результаты показывают, что разработанная система может быть эффективно использована для создания визуально улучшенного контента для цифрового использования.

В работе [165] разработан метод автоматического распознавания цифровой модуляции в сетях связи на основе моделей машинного обучения, оптимизированных с помощью генетических алгоритмов. В данной работе представлен ряд исследований с целью определения эффективности различных алгоритмов идентификации модулированных сигналов (AMR) с широко используемыми цифровыми модуляциями. Многочисленные компьютерные моделирования выполнялись в присутствии аддитивного белого гауссовского шума с отношением сигнал–шум в диапазоне от минус 10 дБ до плюс 30 дБ. Результаты моделирования и сравнения с предыдущими исследованиями показывают, что применение предложенных алгоритмов вместе с выбранными функциями приводит к значительному повышению точности и скорости автоматического определения типов цифровой модуляции при низких значениях соотношениях сигнал/шум (SNR). Кроме того, повышается скорость сходимости моделей.

В работе [166] разработана автоматическая классификация жировой болезни печени на основе контролируемого обучения и генетического алгоритма. Разработанный метод обеспечивает четыре основных вклада: во–первых, достигается классификация изображений печени как нормальных или жировых без фазы сегментации. Во–вторых, по сравнению с предлагаемой работой, набор данных в предыдущих работах был недостаточным. Комбинация из 26 функций является третьим вкладом. На основе предложенных методов извлеченными признаками являются матрица совпадения уровней серого и статистика первого порядка. Четвертый вклад – это классификатор голосования, используемый для определения типа ткани печени. Было проведено изучение классификатора на основе голосования и алгоритма J48 в наборе данных. Полученные значения TP, TN, FP и FN составили 94,28 %, 97,14 %, 5,71 % и 2,85 % соответственно. Достигнутые точность, чувствительность, специфичность и показатель F1 составили 94,28 %, 97,05 %, 94,44 % и 95,64 %, соответственно. Достигнутая точность классификации с использованием классификатора на основе голосования составила 95,71 %, а в случае использования алгоритма J48 – 93,12 %.

1.9.1.3 Автоэнкодеры

Автоэнкодер – это тип нейронной сети, используемый для обучения без учителя, который учится эффективно сжимать и кодировать данные, а затем учится восстанавливать данные обратно из сокращенного представления. Автоэнкодеры специфичны для данных, то

есть они могут сжимать только те данные, которые очень похожи на данные, на которых автоэнкодер уже обучен.

Научно–исследовательская и публикационная активность, связанная с направлением «Автоэнкодеры» за последние 10 лет показана на рисунке 1.9.3 [1, 2].

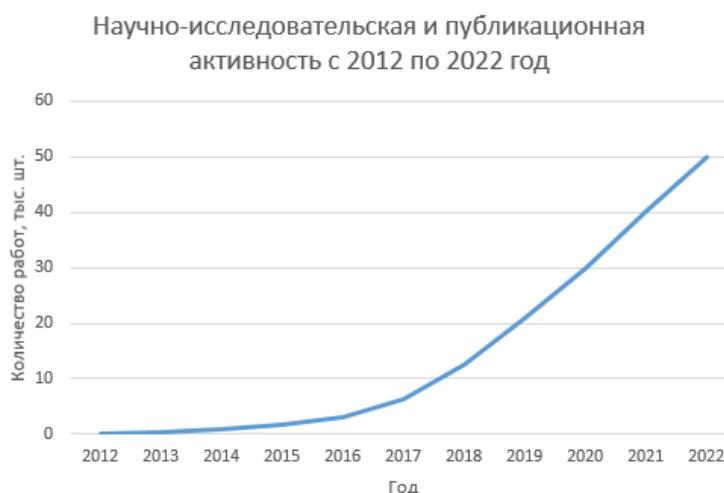


Рисунок 1.9.3 – Научно–исследовательская и публикационная активность, связанная с направлением «Автоэнкодеры» за последние 10 лет

В работе [167] разработан новый метод обработки информации, основанный на наборе автокодеров для неконтролируемого обнаружения неисправностей. Разработанная схема основана на наборе автокодеров различных типов; каждый из них обучается независимо для выполнения задачи обнаружения неисправностей одного класса. Окончательное согласие с нормальностью тестируемого образца достигается в процессе мягкого голосования, при котором учитывается уверенность каждого автокодера в своем решении. Значение каждого отдельного автокодера для принятия окончательного решения извлекается из статистического анализа независимого процесса обучения каждого из них. Результаты моделирования с использованием трех широко используемых наборов данных для обнаружения неисправностей показывают эффективность предложенной модели.

В работе [168] разработан метод оценки надежности энергосистемы с учётом данных на основе встроенных автокодеров с шумоподавлением. Всё большее проникновение возобновляемых источников энергии в энергосистемы непреднамеренно приводит к всплеску числа случайных состояний. Вычисление оптимального сброса нагрузки для всех состояний размножения довольно громоздко. Это считается узким местом в процессе оценки надежности. Для решения этой проблемы предлагается подход, основанный на глубоком обучении, для более эффективной оценки надежности системы с учётом колебаний выработки и нагрузок. В предлагаемой модели потока мощности автокодировщик с шумоподавлением служит многослойной нейронной сетью для процесса глубокого обучения. Численный

результат показывает большое преимущество разработанного метода во времени вычислений при обеспечении высокой точности.

В работе [169] разработан глубокий сверточный подход на основе автоэнкодера для обнаружения аномалий с промышленными двумерными данными, не являющимися изображениями. Доказана эффективность предложенного подхода на реальном примере, связанном с данными оптической эмиссионной спектроскопии в процессе производства полупроводников, обеспечивая удовлетворительную точность классификации.

1.9.2 Оптимизация гиперпараметров

Оптимизация гиперпараметров – это процесс выбора оптимального набора гиперпараметров для модели, которая обеспечивает наилучшую производительность при проверке данных. Гиперпараметры – это параметры модели, которые нельзя определить из данных и которые необходимо задать до обучения модели. Оптимизация гиперпараметров важный шаг в машинном обучении, который позволяет моделям достигать максимальной точности и эффективности обобщения новых данных.

Основными тенденциями развития автоматизированной разработки признаков являются:

- решетчатый поиск,
- случайный поиск,
- оптимизация методов последовательного моделирования.

1.9.2.1 Решетчатый поиск

Решетчатый поиск – это вариант поиска по решетке – распространенный метод оптимизации гиперпараметров. При поиске по сетке пространство поиска определяется как сетка значений гиперпараметров, и каждая комбинация гиперпараметров оценивается для поиска наилучшего набора. Поиск по сетке – простой и исчерпывающий метод, но он может быть дорогостоящим в вычислительном отношении, особенно для многомерных пространств поиска.

Научно–исследовательская и публикационная активность, связанная с направлением «Решетчатый поиск» за последние 10 лет показана на рисунке 1.9.4 [1, 2].

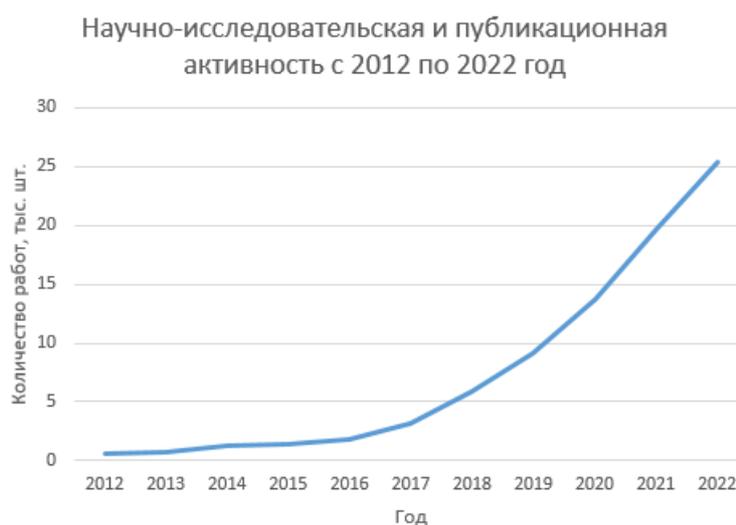


Рисунок 1.9.4 – Научно–исследовательская и публикационная активность, связанная с направлением «Решетчатый поиск» за последние 10 лет

В работе [170] разработан метод основанный на машинном обучении для автоматизации моделирования дыхательных потоков по решетчатому Больцману. За исключением незначительных отклонений на границах раздела между решетчатыми пазухами, сгенерированная сетью поверхность является достаточно точной. Для дальнейшего анализа точности автоматизированного конвейера моделирование потоков проводится с использованием метода тепловой решетки Больцмана в обоих случаях на высокопроизводительной вычислительной системе. Сравнение результатов моделирования дыхательных потоков дает усредненные ошибки менее 1 % для потери давления между входным и выходным отверстиями, а также для температуры на выходе. Таким образом, показано, что трубопровод работает точно, а геометрические отклонения в решетчатых пазухах незначительны.

В работе [171] разработана гетерогенная стратегия на основе глубокого обучения для настройки ответов решетчатых структур. Результаты показывают, что гетерогенные решетчатые структуры более подходят для поглощения энергии, чем монолитные решетчатые структуры с октетной фермой и ромбическим додекаэдром. Прогнозирование методом конечных элементов может быть эффективно и точно воспроизведено искусственной нейронной сетью. Разработанная стратегия расширяет пространство проектирования решетчатых структур и обеспечивает новый подход к проектированию решетчатой структуры с конкретным откликом.

В работе [172] продемонстрировано прогнозирование кристаллической структуры и параметров решетки перовскитовых материалов с помощью различных моделей машинного обучения на основе базовых свойств атома. Точность модели прогнозирования для классификации кристаллической структуры составляет в среднем почти 88 % для GAN, тогда

как для регрессионной модели констант решетки модель «GA-SVR» дает в среднем ~ 95 %, что может быть дополнительно улучшено путем накопления более надежных наборов данных в базе данных. Эти модели машинного обучения могут быть использованы в качестве альтернативного процесса для ускорения разработки новых перовскитных материалов, предоставляя ценную информацию о поведении перовскитных материалов.

1.9.2.2 Случайный поиск

Случайный поиск – это метод, при котором случайные комбинации гиперпараметров выбираются и используются для обучения модели. Используются лучшие случайные комбинации гиперпараметров. Случайный поиск отлично подходит для обнаружения комбинаций гиперпараметров, о которых невозможно догадаться интуитивно.

Научно–исследовательская и публикационная активность, связанная с направлением «Случайный поиск» за последние 10 лет показана на рисунке 1.9.5 [1, 2].

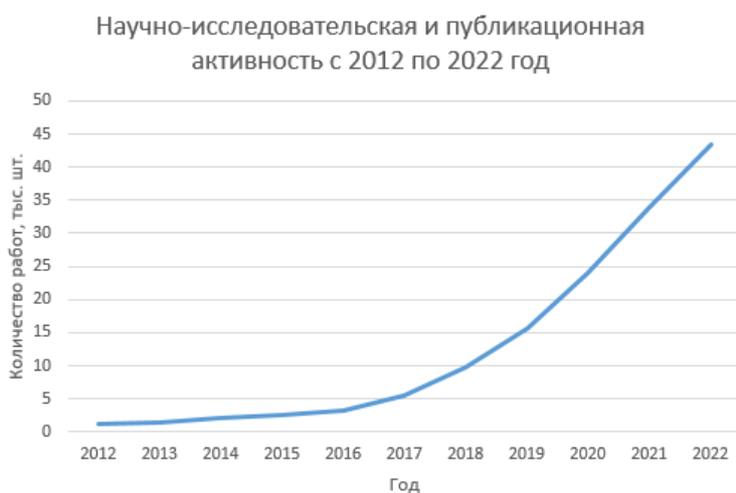


Рисунок 1.9.5 – Научно–исследовательская и публикационная активность, связанная с направлением «Случайный поиск» за последние 10 лет

В работе [173] продемонстрировано сравнение методов машинного обучения для прогнозирования фотоэлектрической мощности на основе численного прогнозирования погоды. Увеличение мировой установленной фотоэлектрической (PV) мощности и прерывистый характер солнечного ресурса подчеркивает важность прогнозирования мощности для интеграции технологии в сеть. В этом исследовании сравниваются 24 модели машинного обучения для детерминистического прогнозирования мощности на сутки вперед на основе численных прогнозов погоды, проверенных на двухлетних наборах данных с 15–минутным разрешением для 16 фотоэлектрических станций в Венгрии. В целом наилучшие прогнозы имеют RMSE на 13,9 % ниже по сравнению с базовым сценарием с использованием линейной регрессии. Более того, прогнозы мощности, основанные только на среднесуточных прогнозах освещенности и углах положения Солнца, имеют RMSE всего на 1,5 % выше, чем

лучший сценарий, что демонстрирует эффективность машинного обучения даже при ограниченной доступности данных. Результаты этого документа могут помочь как исследователям, так и практикам в разработке наилучших методов прогнозирования фотоэлектрической мощности на основе численных прогнозов погоды.

В работе [174] разработан метод, основанный на контексте подхода к обучению с мета–подкреплением для эффективной оптимизации гиперпараметров. В работе используются: переменный контекст, многозадачный объективный метод, метод квадратичного штрафа для повышения производительности агента. В результате модель было оценена на 18 реальных наборах данных, и результаты демонстрируют, что данный подход превосходит другие современные методы оптимизации с точки зрения точности набора тестов и производительности во время выполнения.

В работе [175] рассмотрен автоматизированный эволюционный подход к проектированию составных конвейеров машинного обучения. Эволюционный подход используется для гибкой идентификации структуры конвейера. Для повышения эффективности подхода реализованы дополнительные алгоритмы анализа чувствительности, распыления и настройки гиперпараметров. Кроме того, программная реализация этого подхода представлена в виде фреймворка с открытым исходным кодом. Набор экспериментов проводится для различных наборов данных и задач (классификация, регрессия, прогнозирование временных рядов). Полученные результаты подтверждают правильность и эффективность предложенного подхода при сравнении с современными конкурентами и базовыми решениями.

1.9.2.3 Оптимизация методов последовательного моделирования

Оптимизация методов последовательного моделирования включает поиск наилучшего сочетания линейных и нелинейных моделей для минимизации ошибки регрессии. Для оптимизации последовательной модели могут использоваться различные методы оптимизации. Например, последовательная оптимизация на основе моделей – это популярный подход, который включает подбор одной или нескольких моделей на основе исходного набора наблюдений за производительностью в пространстве конфигураций. Метод подходит для настройки моделей, основанных на статистическом моделировании и методах оптимизации черного ящика.

Научно–исследовательская и публикационная активность, связанная с направлением «Оптимизация методов последовательного моделирования» за последние 10 лет показана на рисунке 1.9.6 [1, 2].

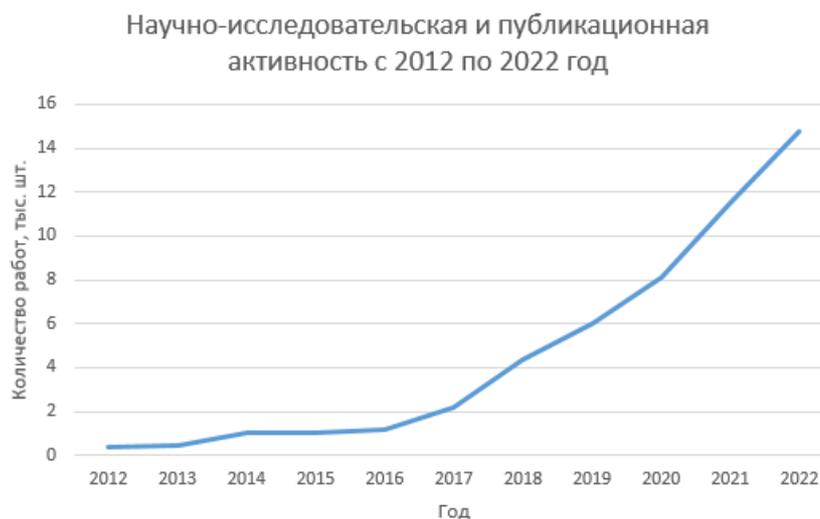


Рисунок 1.9.6 – Научно–исследовательская и публикационная активность, связанная с направлением «Оптимизация методов последовательного моделирования» за последние 10 лет

В работе [176] рассмотрена модель прогнозирования временных рядов сырой нефти на основе сети LSTM с хаотической оптимизацией растворимости газа Генри. Оценка цены сырой нефти, которая рассматривается как важный ресурс экономического развития и стабильности в мире, представляет большой интерес для политиков и участников рынка. Однако хаотические и нелинейные характеристики временных рядов сырой нефти затрудняют оценку цен на сырую нефть с высокой точностью. Чтобы преодолеть эти проблемы, в этом исследовании предлагается новая модель прогнозирования цен на сырую нефть, которая включает в себя долговременную кратковременную память, технические индикаторы, такие как тренд, волатильность и импульс, а также хаотическую оптимизацию растворимости газа Генри. Результаты показывают, что предложенная модель прогнозирования справляется с хаотичностью и нелинейной динамикой.

В работе [177] разработан метод оптимизации энергетического углеродного следа при планировании производства в зависимости от последовательности. В условиях растущей стоимости энергии и экологической политики, вызванной проблемами глобального потепления, это исследование направлено на создание модели графика производства с оптимальным потреблением энергии для минимизации затрат на энергию и сокращения выбросов углерода.

В работе [178] рассмотрена воспроизводимая оптимизация гиперпараметров. Ключевой проблемой в исследованиях машинного обучения является отсутствие воспроизводимости. В работе показывают, какую роль поиск гиперпараметров играет в этой проблеме и как обычные методы поиска гиперпараметров могут привести к большой дисперсии результатов из-за недетерминированного обучения модели во время оптимизации

гиперпараметров. Эмпирические результаты тестов машинного обучения показывают, что при равных вычислениях предлагаемый метод уменьшает вариацию результатов поиска гиперпараметров до 90 %, в то же время приводя к таким же или меньшим средним ошибкам прогнозирования по сравнению со стандартным случайным поиском и байесовской оптимизацией. Более того, структура последовательного тестирования успешно сокращает объём вычислений, сохраняя при этом производительность метода.

1.9.3 Обучение с подкреплением

Обучение с подкреплением (RL, от англ. Reinforcement learning) – это подраздел машинного обучения, в котором алгоритм обучения учится принимать решения, взаимодействуя с окружающей средой. Учащийся получает обратную связь в виде сигнала вознаграждения или сигнала наказания в ответ на свои действия, и его цель состоит в том, чтобы максимизировать общее вознаграждение, которое он получает. Обучение с подкреплением используется для решения проблем, где нет известного алгоритмического решения, и цель состоит в том, чтобы научиться принимать решения на основе опыта.

Основными тенденциями развития обучения с подкреплением являются:

- робототехника,
- автономный транспорт,
- здравоохранение

1.9.3.1 Обучение с подкреплением в робототехнике

Обучение с подкреплением – это область машинного обучения, которая позволяет роботу научиться оптимальному поведению путем проб и ошибок при взаимодействии с окружающей средой. Конечная цель обучения с подкреплением в робототехнике – наделить роботов способностью учиться, улучшать, адаптировать и воспроизводить задачи с динамически меняющимися ограничениями, основанными на реальных задачах предметной области.

Научно–исследовательская и публикационная активность, связанная с направлением «Обучение с подкреплением в робототехнике» за последние 10 лет показана на рисунке 1.9.7 [1, 2].



Рисунок 1.9.7 – Научно–исследовательская и публикационная активность, связанная с направлением «Обучение с подкреплением в робототехнике» за последние 10 лет

В работе [179] рассмотрен метод обучения с подкреплением для планирования роботизированного поточного цеха с изменением времени обработки. На основе данного метода была решена проблема планирования роботизированного производственного цеха, когда два типа деталей обрабатываются на каждом заданном наборе выделенных станков. Один робот, движущийся по неподвижному рельсу, перемещает одну деталь за раз, а время обработки деталей варьируется на машинах в пределах заданного интервала времени. Используется подход обучения с подкреплением, чтобы получить эффективные последовательности задач робота, чтобы минимизировать время ожидания.

В работе [180] рассмотрено глубокое обучение с подкреплением для предотвращения столкновений роботов с самостоятельным вниманием и слиянием датчиков. Датчики 3D LiDAR могут создавать трехмерные облака точек окружающей среды и широко используются в автомобильной навигации; в то время как датчики 2D LiDAR могут обеспечивать облако точек только в плоскости 2D, а затем используются только для навигации роботов небольшой высоты, например, роботов для мытья полов. В работе предлагается новая схема глубокого обучения с подкреплением для навигации роботов, предлагая краткую и эффективную единицу внимания к самосостоянию и доказывая, что применение этой единицы может заменить многоэтапное обучение, достичь лучших результатов и способности к обобщению. Эксперименты как с смоделированными данными, так и с реальным роботом показывают, что разработанный метод может обеспечить эффективное предотвращение столкновений только с использованием недорогого датчика 2D LiDAR и монокулярной камеры.

В работе [181] разработана доказуемая безопасная система манипулирования роботами в среде с людьми с использованием обучения с глубоким подкреплением. В статье используется быстрый анализ доступности людей и манипуляторов, чтобы гарантировать полную остановку манипулятора до того, как человек окажется в пределах его досягаемости.

Предлагаемый метод гарантирует безопасность и значительно повышает производительность глубокого обучения подкреплением, предотвращая столкновения и коллизии.

1.9.3.2 Обучение с подкреплением в автономном транспорте

Обучение с подкреплением используется в автономных транспортных средствах, чтобы научить их принимать решения о вождении на основе окружающей среды. Алгоритмы RL, используемые в автономных транспортных средствах, включают глубокое обучение с подкреплением и мультиагентное обучение с подкреплением. Алгоритмы RL могут помочь автономным транспортным средствам принимать решения в сложных и динамичных условиях, таких как навигация в пробках, обход препятствий и реагирование на непредвиденные события. RL также может помочь повысить энергоэффективность и комфорт вождения автономных транспортных средств.

Научно–исследовательская и публикационная активность, связанная с направлением «Обучение с подкреплением в автономном транспорте» за последние 10 лет показана на рисунке 1.9.8 [1, 2].



Рисунок 1.9.8 – Научно–исследовательская и публикационная активность, связанная с направлением «Обучение с подкреплением в автономном транспорте» за последние 10 лет

В работе [182] разработана система глубокого обучения с подкреплением для проверки безопасности автономных транспортных средств. Одним из тормозящих факторов, препятствующих разработке и развертыванию автономных транспортных средств, являются непомерно высокие экономические и временные затраты, необходимые для проверки их безопасности в естественной среде вождения из-за редкости критических с точки зрения безопасности событий. В статье демонстрируется эффективность разработанного подхода тестами на высокоавтоматизированных транспортных средствах в разных дорожных условиях, например, шоссе и на городских испытательных трассах в среде дополненной реальности, сочетая смоделированные фоновые транспортные средства с физической

дорожной инфраструктурой и реальным автономным испытательным транспортным средством. Результаты показывают, что разработанное решение позволяет ускорить процесс оценки ситуации в 10^3 – 10^5 раз.

В работе [183] разработана система принятия решений автономными транспортными средствами в сценариях смены полосы движения на основе обучения с подкреплением. Предложенные методы были оценены в «CARLA» в двух сценариях (один со статическими препятствиями и один с динамически движущимися транспортными средствами). Результаты показывают, что предлагаемые методы могут генерировать надежные стратегии безопасного вождения и обеспечивать лучшие водительские качества, чем предыдущие методы.

В работе [184] разработан удобный и энергоэффективный контроль скорости автономных транспортных средств на неровных дорожных покрытиях с использованием глубокого обучения с подкреплением. Модель контроля скорости на основе глубокого обучения с подкреплением обучается с использованием реальных данных о неровностях дорожного покрытия в Шанхае. Результаты эксперимента показывают, что комфорт при вертикальной езде, энергоэффективность и эффективность вычислений увеличиваются на 8,22 %, 24,37 % и 94,38 %, соответственно, по сравнению с моделью управления скоростью, основанной на оптимизации.

1.9.3.3 Обучение с подкреплением в здравоохранении

В здравоохранении RL находит широкое применение, так как позволяет обучать компьютерные программы классифицировать и анализировать медицинские данные, улучшая качество медицинских услуг и повышая точность диагностики.

Научно–исследовательская и публикационная активность, связанная с направлением «Обучение с подкреплением в здравоохранение» за последние 10 лет показана на рисунке 1.9.9 [1, 2].



Рисунок 1.9.9 – Научно–исследовательская и публикационная активность, связанная с направлением «Обучение с подкреплением в здравоохранение» за последние 10 лет

В работе [185] рассмотрено использование машинного обучения для решения задач и возможностей здравоохранения. Было выявлено, что важным применением машинного обучения с подкреплением в общественном здравоохранении является выявление и прогнозирование групп населения с высоким риском развития определенных неблагоприятных исходов для здоровья и разработка вмешательств в области общественного здравоохранения, ориентированных на эти группы населения. Различные концепции, связанные с машинным обучением, необходимо интегрировать в медицинскую учебную программу, чтобы медицинские работники могли эффективно направлять и интерпретировать исследования в этой области.

В работе [186] рассмотрен подход к обработке медицинских данных с замкнутым циклом, основанный на обучении с глубоким подкреплением. Основной представленной идеей является разработка виртуального симулятора человеческого тела на основе глубоких нейронных сетей для исследований в области здравоохранения. Симулятор может принимать вмешательства в качестве входных данных и представлять свои скрытые состояния в виде многомерных изображений, которые модуль лечения на основе глубокого обучения с подкреплением может использовать для диагностики скрытых состояний здоровья и выбора подходящих вмешательств для доведения моделируемого тела до целевого состояния. Подход направлен на помощь в принятии решений в области здравоохранения и был подтвержден с помощью экспериментального моделирования.

В работе [187] разработана система поддержки принятия решений при проектировании зданий здравоохранения на основе аргументации на основе конкретных случаев и обучения с подкреплением. Проектирование зданий здравоохранения – это очень сложный процесс, в котором необходимо учитывать большое количество параметров и переменных. Проектирование зданий этого типа и связанных с ними объектов включает в себя большое количество нормативных актов, которые адаптированы к различным странам. Проверка этого регламента очень сложна и обычно требует участия больших команд специализированных инженеров и архитекторов. Предлагаемые рассуждения на основе конкретных случаев и обучение с подкреплением могут анализировать данные о проекте здания (предоставленные в расширенном языке разметки или XML-файле и других совместимых форматах), проверяя и подтверждая правила. Этот подход позволяет сократить штат специализированного и высококвалифицированного персонала, предоставляя отчет с правилами проверки и отслеживая предупреждения и ошибки в применении правил.

1.9.4 Нейронные сети

Нейротехнология относится к любой технологии, которая обеспечивает более глубокое понимание деятельности мозга или нервной системы, или влияет на функцию мозга или

нервной системы. В последние годы нейротехнологии и нейронные сети стали тесно связываться друг с другом и совместно применяться в различных областях таких как медицина, компьютерная наука, физика, биология, химия и др.

Нейронные сети – это тип алгоритма машинного обучения, вдохновленный структурой и функциями человеческого мозга. Они состоят из взаимосвязанных узлов, называемых нейронами, которые работают вместе для обработки и анализа данных.

Основными тенденциями развития обучения с подкреплением являются:

- робототехника,
- предсказание финансовых временных рядов,
- нейроуправление,
- сильный ИИ.

1.9.4.1 Нейротехнологии и нейронные сети в робототехнике

Нейронные сети в робототехнике - это специализированные алгоритмы искусственного интеллекта, основанные на математическом моделировании процессов, происходящих в мозге. Используя такие сети, роботы обрабатывают информацию из окружающей среды, обучаются новым навыкам и выполняют сложные задачи. Нейронные сети приносят в робототехнику: самообучение, универсальность, параллельность задач, распознавание образов, восприятие нечеткой информации и др.

Научно–исследовательская и публикационная активность, связанная с направлением «Нейронные сети в робототехнике» за последние 10 лет показана на рисунке 1.9.10 [1, 2].



Рисунок 1.9.10 – Научно–исследовательская и публикационная активность, связанная с направлением «Нейронные сети в робототехнике» за последние 10 лет

В работе [188] разработана нейронная сети с иерархическим графом переходов для исследования несколькими роботами в неизвестных средах. Мультироботы, проводящие исследование в незнакомых средах по принципу общего к грубому, имеют большое значение

во многих областях применения, таких как поиск и спасение. Для разных этапов задачи роботам необходимо дискриминирующим образом извлекать информацию из окружающей их среды, что улучшает их способность принимать решения. В работе предлагается использовать графовую нейронную сеть Hierarchical-Hops (далее H2GNN), чтобы позволит роботам интегрировать ключевую информацию об окружающей среде, представленную в виде графа, разделяя важность информации с разных уровней относительно роботов на основе механизма множественного вклада внимания. Для повышения эффективности сотрудничества используется мультиагентное обучение с подкреплением, чтобы помочь роботам неявно учиться стратегиям сотрудничества. Также проведены эксперименты для проверки предложенного метода в симуляционной среде, и экспериментальные результаты демонстрируют, что H2GNN значительно улучшает производительность мультироботов в исследовании незнакомых сред.

В работе [189] разработан алгоритм оценки местоположения сельскохозяйственных мобильных роботов на основе слияния мультисенсоров и нейронной сети автоэнкодера. Позиционирование сельскохозяйственных мобильных роботов с высокой точностью является ключевым для выполнения инструкций управления. В работе предлагается улучшенный алгоритм оценки позиции на основе многократной фильтрации и нейронной сети автоэнкодера. Данные многократной фильтрации навигации по спутнику в режиме реального времени, ИМУ и гироскопических датчиков кодера объединяются с помощью расширенного фильтра Калмана. Динамические эксперименты проводились на дорогах, траве и поле. Чтобы проверить устойчивость предложенного алгоритма, аномальные условия работы датчиков тестировались на дороге. Результаты показали, что точность позиционирования улучшается по сравнению с навигацией по спутнику в режиме реального времени во всех трех средах.

В работе [190] разработан алгоритм оптимизации распределенной нейронной сети для совместного обучения с участием нескольких роботов через меш-сеть. Роботы обмениваются весами, чтобы достичь консенсуса. Алгоритм работает с не выпуклыми целевыми функциями, вариативной связью и поточными данными. Результаты экспериментов показали его превосходство над стандартными алгоритмами.

1.9.4.2 Нейротехнологии и нейронные сети в предсказание финансовых временных рядов

Прогнозирование финансовых временных рядов с использованием нейронных сетей относится к использованию моделей глубокого обучения, таких как сверточные нейронные сети (далее CNN, от англ. Convolutional Neural Networks) и сети с долговременной кратковременной памятью, для прогнозирования будущих событий на основе исторических

финансовых данных. Эти модели предназначены для изучения закономерностей в данных и прогнозирования с высокой точностью. Использование нейронных сетей в финансовом прогнозировании может помочь инвесторам и аналитикам принимать более обоснованные решения о покупке, продаже или хранении финансовых активов.

Научно–исследовательская и публикационная активность, связанная с направлением «Предсказание финансовых временных рядов» за последние 10 лет показана на рисунке 1.9.11 [1, 2].



Рисунок 1.9.11 – Научно–исследовательская и публикационная активность, связанная с направлением «Предсказание финансовых временных рядов» за последние 10 лет

В работе [191] предлагается новая методология прогнозирования временных рядов в финансовой отрасли. Авторы предлагают использовать многомодальную графовую нейронную сеть, которая позволяет обрабатывать различные входные данные из разных источников. Для этого создается гетерогенный граф, где узлы представляют источники данных, а ребра - отношения между ними. Для обеспечения интерпретируемости используется механизм двухфазного внимания. Авторы утверждают, что их метод показывает превосходную производительность в прогнозировании финансового рынка и предоставляет инвесторам прибыльный и интерпретируемый вариант принятия инвестиционных решений.

В работе [192] обсуждаются методы прогнозирования временных рядов с использованием обобщенных регрессионных нейронных сетей. Авторы приводят стратегии для принятия ключевых модельных решений в прогнозировании временных рядов описывает оригинальные техники работы с трендами и сезонными закономерностями. Был представлен пакет R, который включает в себя лучшие подходы и трансформации для прогнозирования временных рядов с помощью обобщенных регрессионных нейронных сетей. Авторы демонстрируют эффективное использование внутренних свойств нейронных сетей для создания быстрых, высокоточных прогнозов автоматически.

В работе [193] было предложено новое решение проблем долгосрочных зависимостей и хаотических свойств при прогнозировании цен на акции с помощью глубоких нейронных сетей. Авторы предложили преобразовывать временные ряды в сложные сети, конвертируя последовательности цен на рынке в графы, и использовать полученные данные для модели прогнозирования. В итоге, в проводимых имитационных торговых операциях модель показала наивысшую накопленную прибыль, а в проведенном анализе результатов модель продемонстрировала лучшие результаты по сравнению с рядом передовых методов.

1.9.4.3 Нейроуправление

Нейроуправление и нейронные сети имеют тесную взаимосвязь, так как нейронные сети являются основой для создания систем нейроуправления. Системы нейроуправления используются для управления различными механизмами и устройствами. Они работают на основе нейронных сетей, которые обрабатывают входные данные и генерируют соответствующие выходные сигналы.

Научно–исследовательская и публикационная активность, связанная с направлением «Нейроуправление» за последние 10 лет показана на рисунке 1.9.12 [1, 2].



Рисунок 1.9.12 – Научно–исследовательская и публикационная активность, связанная с направлением «Нейроуправление» за последние 10 лет

В работе [194] разработан новый алгоритм обучения для нейроконтроля нелинейных систем с дискретным временем. Физические системы являются нелинейными, что затрудняет их управление. Традиционные методы управления такими системами используют линейное приближение первого порядка, а нелинейные методы требуют аналитических моделей. Искусственные нейронные сети используются для управления нелинейными системами, и был предложен алгоритм обратного распространения ошибки для обучения рекуррентных сетей в статье Нарендра и Партхасарати в 1990 году. Однако этот алгоритм имеет медленную сходимость, поэтому был разработан новый алгоритм последовательного обучения с лучшими

свойствами сходимости. Данный алгоритм распространяется на сети с одним скрытым слоем, однако авторы предлагают расширить его применимость на сети с двумя скрытыми слоями. Для этого авторы предлагают использовать метод обратного распространения ошибки. Этот новый алгоритм обладает лучшими свойствами сходимости и более эффективен в управлении нелинейными системами.

В работе [195] рассматривается проблема устойчивого нейроконтроля неизвестных нелинейных систем с асимметричными ограничениями ввода. Для ее решения используется функция со скидкой для номинальных систем, которая преобразует задачу устойчивого ограниченного контроля в проблему оптимального контроля с нелинейным ограничением. Для решения этой проблемы используется архитектура субъект-критик, в которой две нейронные сети используются для приближения политики управления и функции стоимости соответственно. При этом веса субъекты и критики определяются методом наименьших квадратов и методом Монте-Карло. В результате проведенных моделирований на нелинейной системе показана эффективность применения предложенного подхода.

В работе [196] описывается метод вычисления функции Ляпунова управления для нелинейных динамических систем на основе стратегии обучения скоростной нейроуправляемости. Предложенный метод использует глубокую нейронную сеть для генерации управляющих функций, поддержанных теорией устойчивости Ляпунова. Также производится оценка области притяжения для продвинутого анализа устойчивости системы. Были проведены эксперименты для сравнения с другими методами, и результаты показали, что предложенный метод дает лучшие решения для нелинейных систем в проектировании стабильных элементов управления без линейных аппроксимаций и в присутствии помех.

1.9.5 Сильный искусственный интеллект

Сильный искусственный интеллект, также известный как искусственный общий интеллект (далее AGI, от англ. Artificial general intelligence), относится к конструкции искусственного интеллекта, которая обладает умственными способностями и функциями, имитирующими человеческий мозг. AGI предназначен для выполнения любой задачи, на которую способен человек. Он представляет собой представление обобщенных когнитивных способностей человека в программном обеспечении, позволяющем системе AGI находить решения незнакомых задач.

Научно–исследовательская и публикационная активность, связанная с направлением «Сильный искусственный интеллект» за последние 10 лет показана на рисунке 1.10.5 [1, 2].



Рисунок 1.9.13 – Научно–исследовательская и публикационная активность, связанная с направлением «Сильный искусственный интеллект» за последние 10 лет

В работе [197] рассмотрен эволюционный подход к нахождению уравнений в частных производных на основе данных, применяемый для моделирования физических процессов. В работе предложен метод улучшения качества результирующего решения путем внедрения более надежного метода оценки качества предлагаемых структур, основанного на включении автоматизированного алгоритма решения уравнений в частных производных компаниях. Это может обеспечить более точное моделирование процессов и позволить обойти сложности, связанные с выводом уравнений в частных производных из моделируемой системы. Однако введение нового метода требует дополнительных вычислительных ресурсов. В работе были проведены эксперименты на синтетических данных, полученных в результате решения уравнения Кортевега-де Фриза, которые подтверждают целесообразность использования нового метода.

В работе [198] авторы предлагают вероятностное обобщение известной концепции Strong Backdoor Set (SBS), применимой к проблеме булевой выполнимости (SAT). Набор булевых переменных называется ρ -черным по доле не менее ρ возможных назначений переменных, если их назначение удовлетворяет булевой формуле в конъюнктивной нормальной форме (CNF) и приводит к полиномиально разрешимым формулам. Авторы определяют (ϵ, δ) -VSB как множество B , для которого параметр ρ отклоняется от 1 не более чем на ϵ с вероятностью не менее $1 - \delta$. Обсуждаются проблемы нахождения минимума VRMB и минимума (ϵ, δ) -VRMB и предлагается новый вероятностный алгоритм решения последней задачи. Они также предлагают метод метаэвристической оптимизации, основанный на штрафных функциях, для практического применения поиска минимума (ϵ, δ) -SBS. Вычислительные эксперименты авторов показывают, что использование (ϵ, δ) -SBS, найденного предложенным алгоритмом, ускоряет решение тестовых задач, связанных с

проверкой эквивалентности и сложными комбинаторными тестами, по сравнению с современными SAT-решателями.

В работе [199] был предложен асинхронный параллельный эволюционный алгоритм для оптимизации задач серого ящика с разбиением на независимые вычисления. Один из таких примеров - поиск бэкдоров в задаче булевой выполнимости. Эксперименты показывают, что предложенный алгоритм ускоряет процесс оптимизации и эффективно использует вычислительные ресурсы.

В работе [200] рассматриваются генетические алгоритмы. Многие задачи дискретной оптимизации имеют плато, что приводит к трудностям в их решении эволюционными алгоритмами. Однако, высокая частота мутаций может помочь совершить прыжок с плато к более лучшей точке, но в целом алгоритмы проводят случайные блуждания по плато. Наше исследование показало, что генетический алгоритм $(1+(\lambda,\lambda))$ способен решать задачу вершинного покрытия более эффективно, чем более простые эволюционные алгоритмы, основанные только на мутациях. Важную роль при решении таких задач играет структура задачи и способ использования кроссоверов. Таким образом, было обнаружено, что дрейф в сторону точек, где найти следующее улучшение проще, существует только для определенного класса экземпляров и подмножества точек поиска, но все же отвечает за повышение производительности в гораздо большем диапазоне случаев.

В работе [201] исследовалась возможность применения мультиагентных систем для обеспечения безопасности и защиты нефтепроводов в контексте пространственного граничного контроля для консенсуса МАС дробного порядка с членом реакции-диффузии. Были изучены два метода согласованного пограничного контроля, основанные на граничных условиях Неймана, и представлены согласованные критерии для МАС в терминах LMI, используя неравенства дробного порядка. Приведены два численных примера, которые демонстрируют эффективность предложенных теоретических результатов в среде многоагентного моделирования. В целом, данная статья подчеркивает значимость МАС в области безопасности нефтегазовой отрасли и предлагает новый подход для ее обеспечения через использование пространственного граничного контроля.

В работе [202] рассмотрены модели корреляции событий безопасности с точки зрения их представления в системах мониторинга на основе ИИ, такие как модели на основе правил, семантики, графики и машинного обучения. Также рассмотрены основные направления текущих исследований в области корреляции событий безопасности на основе ИИ и методы, используемые для корреляции одиночных событий и их последовательностей в сценариях атак, а также перспективы развития гибридных корреляционных моделей. В заключении указаны существующие проблемы в сфере и возможные пути их преодоления. Таким образом,

данный обзор представляет собой систематизацию моделей корреляции событий безопасности и обзор современных исследований в этой области.

В работе [203] описывается разработка роботизированных захватов для использования в роботах, протезах и ортезах. Эти захваты должны быть настраиваемыми и соответствовать широкому кругу требований, включая адаптивность, ловкость, высокое отношение полезной нагрузки к весу, прочность, эстетику, компактность и легкость. В разработанном пятипальцевом захвате с 14 степенями свободы применены встроенные пассивные звенья переменной длины и эластичные элементы во входных суставах, что обеспечивает адаптивность пальцев. Разработанная рука с четырьмя моторами может выполнять как точные, так и силовые захваты различных объектов. Документ содержит описание связей для пальцев, обзор дизайна руки, стратегию управления и результаты испытаний физического прототипа.

В работе [204] было рассмотрено, что управление дорожным движением. Была разработана эффективная модель оптимизации транспортной сети, которая с помощью датчиков собирает информацию о транспортном потоке в режиме реального времени и оптимизирует работу всей транспортной системы. Многоагентный подход с распределенным алгоритмом оптимизации был использован для решения этой проблемы. Цель исследования заключается в минимизации общей стоимости трафика сети путем настройки оптимального равновесного потока трафика и это может быть важным ориентиром для решения проблемы оптимизации заторов в городах.

В работе [205] авторы исследования представляют основу для обучения роботизированных систем с использованием генеративной модели сред. Они предлагают неявный метод указания априорных политик, который обновляется с помощью данных о реальных средах. Используя теорию РАС-Байесовского обобщения, они могут минимизировать верхнюю границу ожидаемых затрат в новых средах. Авторы демонстрируют свой подход на квадрокоптерной навигации и задаче захвата. Результаты работы подтверждают способность их подхода получать более сильные гарантии обобщения в сравнении с предыдущими работами. Они также проводят аппаратные эксперименты для проверки границ метода на задаче захвата.

В работе [206] описывается метод автоматической классификации типов сварных швов, распознавания точек прихватки и определения ROI сварки на основе модифицированного YOLOv5. Требования к классификации типов сварных швов, распознаванию места прихватки и определению ROI преобразованы в единую задачу локализации цели для повышения скорости вывода. В результате экспериментов было достигнуто снижение смещения центрального компонента определенной области интереса сварки с 2,38 пикселя до 2,18

пикселя, уменьшение количества ложных срабатываний при сварке угловых и разделочных швов, а также улучшение времени расчета с 180 мс до 48 мс при повышении полноты с 0,96 до 0,98.

В работе [207] подводятся итоги использования искусственного интеллекта в медицинских системах. ИИ показал чрезвычайно высокие результаты в решении некоторых задач, превосходя человека. Однако, для успешного использования ИИ в медицине, необходимо создание надежных систем, способных справляться с неточной и неверной информацией и объяснять результат эксперту. Для достижения этой цели, важно объединить усилия в трех областях: сложные сети и их выводы, графические причинно-следственные модели и контрфактуалы, а так методы проверки и объяснения. Важно учитывать этические и правовые аспекты в использовании ИИ в медицине. Комплексное и интегративное слияние информации позволит помочь объединить эти три области и сделать будущее медицинского ИИ более надежным и этически ответственным.

В работе [208] продемонстрирован медицинский искусственный интеллект, который позволяет системам достигать производительности на уровне клинических экспертов. Однако при оценке в условиях, отличных от условий обучения, эти системы часто демонстрируют субоптимальную производительность. Один из возможных способов решения этой проблемы - разработка отдельных систем для каждой клинической ситуации, но это требует много времени и дорогостоящей аннотации данных. REMEDIS представляет собой унифицированную стратегию обучения представлению для эффективной обработки данных медицинских изображений с использованием комбинации крупномасштабного контролируемого трансферного обучения с самоконтролируемым обучением. REMEDIS позволяет значительно улучшить производительность, а также эффективно обобщать и обрабатывать данные ИИ для медицинских изображений с использованием относительно небольшого количества данных переобучения. REMEDIS может значительно ускорить жизненный цикл разработки ИИ для медицинских изображений, что позволяет ИИ для медицинских изображений оказывать широкое влияние в медицинской индустрии.

В работе [209] была рассмотрена проблема изменения экземпляров задачи в процессе оптимизации. Было выяснено, что использование лучших решений предыдущих экземпляров может не привести к оптимальному результату, так как эволюционные алгоритмы могут заменять хорошие решения на более плохие. Для решения этой проблемы был предложен модифицированный алгоритм повторной оптимизации, который выполняет интерполяцию между жадным поиском вокруг предыдущего лучшего и текущего лучшего решения. Это позволило достичь лучших результатов, чем использование простого лучшего решения предыдущего экземпляра или полная перезапуск каждого экземпляра. Эксперименты были

проведены на задаче LeadingOnes с различной частотой изменений и факторами возмущения.

В работе [210] проведено исследование, которое показывает, что распознавание эмоций на основе визуальных эффектов является сложной задачей из-за отсутствия способности к обобщению. Однако, введение визуального кросс-корпусного исследования и сквозной структуры распознавания эмоций на основе визуальных эффектов, позволяет выполнять эту задачу с достаточной точностью. Магистральная модель, предложенная в этом исследовании, продемонстрировала высокую способность к обобщению и достигла точности 66,4 % в наборе данных AffectNet, превзойдя все современные результаты. Модель CNN-LSTM также продемонстрировала достойную эффективность на динамических визуальных наборах данных во время кросс-корпусных экспериментов.

1.9.6 Вывод

Основными тенденциями развития автоматического машинного обучения являются глубокое обучение, генетические алгоритмы, автоэнкодеры в автоматизированной генерации признаков, решетчатый поиск, случайный поиск, оптимизация методов последовательного моделирования в оптимизации гиперпараметров, обучение с подкреплением, робототехника, предсказание финансовых рядов с использованием нейронных сетей, нейроуправление и сильный искусственный интеллект.

1.10 Системы обработки больших объёмов данных

Вопрос автоматического машинного обучения вынесен в отдельный подраздел, так как связь между большими данными и искусственным интеллектом имеет основополагающее значение для преобразования исходных данных в информацию, которая может прогнозировать и направлять нас к более эффективным решениям.

«Большие данные» – это новая концепция, которая до сих пор не может получить правильное и стандартное определение. Чтобы придать этой концепции какой-то осязаемый смысл, эксперты разделили ее на 3 основных релевантных сегмента:

- разнообразие – вся информация, хранящаяся в больших данных, чрезвычайно разнообразна, и каждая из них поступает из очень конкретного источника;
- скорость – для архивирования, сохранения и каталогизации такого большого объёма информации требуется чрезвычайная скорость;
- объём – необходимы технологические решения, позволяющие управлять всеми массивными размерами больших данных.

Системы обработки больших данных – это комплекс программных и аппаратных средств, предназначенных для обработки, хранения и анализа больших объёмов данных. Данная технология обеспечивает возможность обработки данных, которые не могут быть

обработаны с помощью традиционных методов и инструментов, включая многопоточные, распределенные и параллельные алгоритмы обработки данных.

Агрегирование огромных массивов данных, начальная точка которых измеряется в терабайтах (1024 гигабайта), петабайтах (1024 терабайта) и даже зеттабайтах (миллиарды терабайт), без применения алгоритмов машинного обучения превращает большие данные из капитала в бремя для бизнеса.

1.10.1 Обработка больших объёмов данных в режиме реального времени

Обработка больших объёмов данных в режиме реального времени – это процесс сбора, обработки и анализа больших объёмов данных, который осуществляется без задержек в обработке и передаче данных.

Основными стимуляторами развития обработки больших объёмов данных в режиме реального времени являются:

- электронная коммерция, финансовые услуги и банковское дело,
- здравоохранение,
- промышленность.

1.10.1.1 Электронная коммерция, финансовые услуги и банковское дело

Обработка больших данных в режиме реального времени имеет решающее значение в различных отраслях, таких как электронная коммерция, финансовые услуги и банковское дело. Использование обработки данных в режиме реального времени позволяет обнаруживать мошенничество, отслеживать ежедневные транзакции и выявлять потенциальные возможности роста.

Научно–исследовательская и публикационная активность, связанная с направлением «Электронная коммерция, финансовые услуги и банковское дело» за последние 10 лет показана на рисунке 1.10.1 [1, 2].



Рисунок 1.10.1 – Научно–исследовательская и публикационная активность, связанная с

направлением «Электронная коммерция, финансовые услуги и банковское дело» за последние 10 лет

В работе [211] рассмотрена передача данных трансграничной электронной коммерции в режиме реального времени на основе больших данных с учётом сложности данных. Функция уровня терминального устройства состоит в том, чтобы обеспечить согласованный интерфейс доступа для вызова ядра, чтобы ядру не приходилось учитывать тип конкретного терминала при работе с последовательным портом. Результаты показывают, что данный метод способен повысить эффективность передачи данных.

В работе [212] разработана система управления финансовыми услугами цепочки поставок на основе обмена данными IoT по блочной цепочке и граничных вычислений. Работа основана на теоретических исследованиях, в ней анализируется финансирование цепочки поставок и технология цепочки блоков. В сочетании с текущей конкретной ситуацией цепочки блоков в финансировании цепочки поставок анализируется система управления, движение денежных средств в цепочке поставок и система контроля рисков. Все участники финансирования цепочки поставок оптимизируют систему контроля рисков финансирования цепочки поставок, одновременно снижая затраты бизнеса и повышая корпоративную эффективность, что значительно снижает риски всех сторон финансирования цепочки поставок. Среда блокчейна «ИОТА», основанная на общих данных и расширенной обработке данных, имеет очень мощное теоретическое и практическое значение для содействия развитию коммерческих банков и предприятий.

В работе [213] рассмотрен прогнозный анализ кредитного риска в режиме реального времени с помощью интеллектуального мониторинга и техники машинного обучения. В данном исследовании была использована технология «Random forest» для создания модели прогнозирования дефолта на основе исторических данных о кредитах из банков и других финансовых институтов. Эта техника превосходит другие классификационные алгоритмы, такие как деревья решений и логистическая регрессия, по точности и скорости воспоминания в соответствии с экспериментальными данными. Кроме того, с помощью техники машинного обучения случайного леса были определены ранжирования функций набора данных. Финансовый сектор может лучше предсказывать риски кредитования и контролировать предварительные риски посредством получения характеристик, которые имеют значительное влияние на риск займа.

1.10.1.2 Здравоохранение

Системы здравоохранения генерируют огромное количество данных из различных источников, таких как электронные медицинские карты, медицинские изображения, геномика и Интернет вещей. Наука о данных играет важную роль в управлении, анализе и усвоении этих

данных для получения фактических результатов. Обработка данных в режиме реального времени стала важнейшим элементом управления больницами, хирургическими отделениями, и др.

Научно–исследовательская и публикационная активность, связанная с направлением «Здравоохранение» за последние 10 лет показана на рисунке 1.10.2 [1, 2].



Рисунок 1.10.2 – Научно–исследовательская и публикационная активность, связанная с направлением «Здравоохранение» за последние 10 лет

В работе [214] рассмотрена аналитика больших данных на основе IoT в здравоохранении. Разработка устройств IoT привела к генерации большого объёма данных, и необходимо создать интегрированную систему аналитики больших данных, основанную на IoT и облачных технологиях, которая позволит улучшить производительность системы. Эта работа представляет собой систему, которая использует IoT и облачные технологии для хранения данных и анализа информации, собираемой от датчиков в реальном времени. Эта система была успешно протестирована в мониторинге здоровья пациентов и показала положительную динамику в их обследовании благодаря автоматическому наблюдению за какими–либо изменениями в их здоровье в режиме реального времени.

В работе [215] разработан метод федеративного обучения с расширенными возможностями обработки медицинских данных в режиме реального времени для интеллектуального здравоохранения. В работе рассматриваются проблемы использования технологии глубокого обучения в медицинской сфере связанные со сбором медицинских данных. Для решения этой проблемы авторы предлагают метод обработки медицинских данных в реальном времени на основе федеративного обучения, который предполагает разделение процесса на стадии модели и примера. Экспериментальные результаты показывают, что этот метод эффективно может обучать модель диагностики заболеваний на основе непрерывного потока медицинских данных в режиме реального времени и смягчить проблему катастрофического забывания новой модели.

В работе [216] разработана система анализа больших данных на основе Building Information Modeling (далее BIM) для управления медицинским учреждением. В данном исследовании рассмотрено использование BIM для оптимизации Facility Management (управление коммунальными службами) в здравоохранении. BIM обычно не применяется из-за своей статичности в то время как управление коммунальными службами требует динамической обработки информации. Предложенная система, созданная на основе Big Data Analytic (далее BDA), BIM и Not only SQL (далее NoSQL) баз данных, позволяет осуществлять коммунальное управление с возможностью запроса и визуализации ключевых показателей. Эффективность системы проверена в реальной ситуации и подтверждена профессионалами в области здравоохранения.

1.10.1.3 Промышленность

Обработка больших объёмов данных в режиме реального времени имеет незаменимое значение в производственной отрасли, поскольку позволяет собирать, анализировать и использовать данные в режиме реального времени, что приводит к улучшению процесса принятия решений, повышению эффективности и сокращению времени простоя. Обработка в реальном времени включает в себя обработку данных по мере их создания с минимальной задержкой и создание автоматических ответов или отчетов в реальном времени.

Научно–исследовательская и публикационная активность, связанная с направлением «Промышленность» за последние 10 лет показана на рисунке 1.10.3 [1, 2].



Рисунок 1.10.3 – Научно–исследовательская и публикационная активность, связанная с направлением «Промышленность» за последние 10 лет

В работе [217] рассмотрено устойчивое интеллектуальное производство на основе больших данных. Устойчивое интеллектуальное производство снижает затраты, повышает производительность и гибкость, улучшает качество и устойчивость в контексте управления жизненным циклом продукта. В данном исследовании рассмотрены технологии «цифровой двойник» и большие данные для создания механизма управления информационно–

энергетическими системами в энергоемких отраслях. Исследование также демонстрирует эффективность стратегии с точки зрения энергосбережения и снижения затрат на охрану окружающей среды в двух случаях в Южном и Северном Китае. Полученные результаты свидетельствуют об эффективности применения стратегии и дают рекомендации для управления энергоемкими отраслями в Китае.

В работе [218] рассмотрено применение интегрированных технологий блокчейна и больших данных для улучшения отслеживания цепочки поставок и обмена информацией в текстильном секторе. Сегодня текстильная промышленность становится все более эффективной, но существуют серьезные проблемы в управлении цепочкой поставок, такие как подделка, плохая прослеживаемость, задержки и отсутствие обмена информацией в реальном времени. Технология блокчейн может решить эти проблемы благодаря своим важным функциям, таким как децентрализация, прозрачность и неизменность. В данном исследовании предлагается новая система прослеживаемости цепочки поставок текстильной промышленности на основе блокчейн, которая может обеспечивать платформу для обмена информацией и прозрачности. Это поможет покупателям принимать информированные решения при покупке и поддерживать компании, которые следуют этическим принципам. Также это поможет улучшить отношения между участниками цепочки поставок и снизить риски и расходы на отзывы товара, подделки и недобросовестный труд. Однако технология блокчейн все еще имеет свои недостатки, в том числе масштабируемость

В работе [219] представлена система управления возобновляемыми источниками энергии в интеллектуальных сетях с использованием анализа больших данных и машинного обучения. Была разработана система для потенциальной реализации аналитики больших данных для умных сетей и энергетических утилит на возобновляемых источниках энергии. Предложен пятиступенчатый подход для предсказания стабильности умной сети с использованием методов машинного обучения. Используя данные из децентрализованной системы умной сети, удалось предсказать стабильность системы с точностью до 96 % при использовании регрессионной модели. Работа имеет ограничения в размере набора данных, однако облачные вычисления и анализ реального времени были пригодны для рамки аналитики больших данных.

1.10.2 Облачная обработка больших объёмов данных

Облачная обработка больших объёмов данных – это использование облачных вычислительных ресурсов и технологий для сбора, хранения, обработки и анализа больших объёмов данных, которые обычно невозможно обработать на обычных вычислительных устройствах. Облачная обработка больших объёмов данных обеспечивает высокую производительность, гибкость, масштабируемость и доступность.

Основными тенденциями развития облачной обработки больших объёмов данных являются:

- облачная обработка больших объёмов данных в здравоохранение и медико–биологических науках;
- облачная обработка больших объёмов данных в розничной торговле и электронной коммерции;
- облачная обработка больших объёмов данных в промышленной Интернет вещей.

1.10.2.1 Облачная обработка больших объёмов данных в здравоохранении и медико–биологических науках

Облачные вычисления – это перспективная технология, способная предоставлять по запросу услуги по хранению, обработке и анализу данных, что особенно актуально для учреждений со сложившейся локальной инфраструктурой – в здравоохранении и биомедицинских науках из–за быстрого роста электронных данных, ориентированных на пациента, и создания больших данных. Облачные вычисления предлагают пользователям доступ с оплатой по мере использования к таким услугам, как аппаратная инфраструктура, платформы и программное обеспечение для решения общих биомедицинских вычислительных задач.

Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших объёмов данных в здравоохранение и медико–биологических науках» за последние 10 лет показана на рисунке 1.10.4 [1, 2].



Рисунок 1.10.4 – Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших объёмов данных в здравоохранение и медико–биологических науках» за последние 10 лет

В работе [220] разработана делегированная архитектура «Quantum Cloud» на основе

блокчейна для защиты больших медицинских данных. Представлено решение проблемы квантовых вычислений в системах здравоохранения – квантовое облачное решение как сервис. Новизна данного решения заключается в использовании квантовых терминальных машин и технологии блокчейн для повышения возможности и безопасности предложенной архитектуры. Экспериментальные результаты доказывают возможность использования данной архитектуры и абсолютную безопасность внедренной системы шифрования «Q-OTR».

В работе [221] рассматриваются биомедицинские облачные платформы для факторного анализ информационного риска (далее FAIR, от англ. Factor analysis of information risk). Биомедицинское исследовательское сообщество инвестирует огромные средства в биомедицинские облачные платформы. Облачные вычисления обещают решить проблемы с «большими данными» и обеспечить воспроизводимость в биологии. Однако, несмотря на все преимущества, облачные платформы, как таковые, не обеспечивают автоматическую исполнительную FAIR. В данной статье утверждается, что эти трудности являются системными и возникают из-за поддержки инсценированных платформ и репозитория данных вместо интероперабельных микросервисов. Также в работе заявляется, что акцентирование на модульности и интероперабельности привело к более мощной «Unix» подобной экосистеме веб-сервисов для биомедицинского анализа и извлечения данных.

В работе [222] рассматриваются облачные приложения IoT в биомедицине. В данном тексте рассматривается технология IoT, которая позволяет связывать биомедицинские и имплантируемые медицинские устройства через сеть и передавать данные по сети без участия человека. При этом обработка данных, обеспечение безопасности и другие процессы IoT осуществляются с помощью облачных вычислений. Однако при передаче биомедицинских данных по сети возможны нарушения безопасности, поэтому для их защиты предлагается интеграция IoT и облачных технологий. Для успешной интеграции IoT и облачных технологий в работе предлагается ввести надежные и безопасные методы валидации и IoT-фреймворк.

1.10.2.2 Облачная обработка больших объёмов данных в розничной торговле и электронной коммерции

Облачная обработка больших данных оказывает значительное влияние на электронную коммерцию и розничную торговлю. Обработка больших данных в среде облачных вычислений становится все более важной из-за сложности управления и эффективного использования соответствующих данных в традиционной ИТ-инфраструктуре. Облачные вычисления и технологии больших данных создали новую бизнес-модель и принесли людям большое удобство. Облачные вычисления можно использовать для унифицированного управления ресурсами и оптимизации ресурсов данных, подключенных к Интернету. Массивные высококачественные данные могут быть использованы для оптимизации

маркетинговых услуг. Точный маркетинг, основанный на облачных вычислениях, заключается в интеграции различных сетевых ресурсов, полученных из нескольких каналов, для достижения более широкого спектра облачных платформ.

Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших объёмов данных в розничной торговле и электронной коммерции» за последние 10 лет показана на рисунке 1.10.5 [1, 2].



Рисунок 1.10.5 – Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших объёмов данных в розничной торговле и электронной коммерции» за последние 10 лет

В работе [223] рассмотрено влияние аналитики больших данных и облачных вычислений на индустрию электронной коммерции. Использование аналитики больших данных помогает оперативно находить недостатки и преимущества принятых решений и делать более обоснованные бизнес–решения. Также аналитика больших данных при наличии достаточного количества данных о пользователе позволяет безграничном масштабе генерировать персональные предложения, от которых обычному пользователю тяжело отказаться. Как показывает опыт внедрения, одним из непредвиденных результатов использования технологии при покупках в Интернете является повышенная склонность к развитию покупательной зависимости. На 2023 год конкуренция на современном рынке становится все жёстче, большинство компании вынуждены инвестировать в дорогостоящее оборудование и нанимать специалистов по обработке больших данных. Также существует тенденция перехода на облачную обработку больших объёмов данных среди малого и среднего бизнеса.

В работе [224] предоставлено исследование системы интеллектуального анализа электронной торговли на основе компьютерных больших данных и облачных вычислений. Данная статья исследует и разрабатывает прототип системы электронной коммерции для

международных покупок. Для создания системы использовались инструменты разработки Eclipse + Tomcat + MySQL и фреймворки Java spring + spring + mybatis для управления товарами, контентом, заказами, пользователями, логистикой и платежами. Система также позволяет осуществлять статистический анализ, поиск товаров, покупки и другие функции, обеспечивая близкий контакт между потребителями, продавцами и логистикой, предоставляя удобную платформу для международных покупок.

В работе [225] рассмотрена безопасность сети электронной коммерции на основе больших данных в среде облачных вычислений. Авторы сосредоточены на методах оценки рисков в областях физической безопасности, на доступе к данным и на их резервном копировании. Облачные вычисления позволяют решать проблемы неравномерного распределения данных и обеспечивать балансировку нагрузки. В статье приводятся рекомендации по улучшению безопасности и доступности системы. В результате всестороннего исследования авторы считают, что комплексный риск–фактор данной системы составляет 4,4. С учётом этого, статья предлагает соответствующие меры по снижению рисков.

1.10.2.3 Облачная обработка больших объёмов данных в промышленном Интернете вещей

Промышленный Интернет вещей (далее IIoT, от англ. Industrial Internet of Things) – это подмножество Интернета вещей, которое относится к взаимосвязи датчиков, инструментов и других устройств, используемых в промышленных приложениях, таких как производство и управление энергопотреблением.

Облачная обработка больших данных в производстве и промышленном Интернете вещей очень актуальна, когда объём данных, генерируемых производственными системами, быстро растёт. Облачные вычисления и аналитика больших данных – это основные технологии, которые могут расширить возможности интеллектуальных производственных систем.

Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших объёмов данных в промышленной Интернет вещей» за последние 10 лет показана на рисунке 1.10.6 [1, 2].



Рисунок 1.10.6 – Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших объёмов данных в промышленной Интернет вещей» за последние 10 лет

В работе [226] разработан сервисно–ориентированный промышленный шлюз Интернета вещей для облачного производства. Облачное производство представляет собой сервисно–ориентированную парадигму производства, которая обеспечивает доступ к различным настраиваемым производственным услугам в облаке в любое время по требованию. Эффективное получение, обмен, хранение, запрос и анализ данных производственного оборудования на уровне производственной площадки остаются значительной проблемой, которая затрудняет развитие систем облачного производства. В данной статье исследуется внедрение новых технологий промышленного интернета вещей в системе облачного производства, чтобы решить данную проблему. В работе проверили жизнеспособность и преимущества предложенного подхода практической реализацией PoT–шлюзов на 3D–принтере и станке. Исследование показывает, что специально разработанные сервисно–ориентированные схемы данных, захватывающие необходимую информацию для принятия высокоуровневых решений облачного производства через технологии PnP PoT, являются хорошим решением для связи производственного оборудования на уровне производственной площадки с облачной платформой производства.

В работе [227] продемонстрирована защищенная сетевая модель от бот–атак в периферийном промышленном Интернете вещей. Экспериментальные результаты демонстрируют многообещающие показатели эффективности в терминах точности, прецизионности, полноты, ошибочных рассуждений, F–меры и коэффициента корреляции Мэтьюса против известных и неизвестных бот–атак.

В работе [228] рассмотрена персонализированная разгрузка задач с учётом конфиденциальности для промышленного Интернета вещей с поддержкой пограничного

облака в автоматизированном производстве. Для обеспечения желаемого уровня качества обслуживания при передаче задач планирования расписания в сервер пограничных вычислений точная информация об интеракции между производственной линией и сервером пограничных вычислений должна быть загружена на облачную платформу, что ставит под угрозу конфиденциальность данных. Существующие работы большей частью предполагают, что все данные интеракции, то есть решения о передаче подзадачи в задаче планирования расписания, имеют одинаковый уровень конфиденциальности, что не удовлетворяет различным требованиям конфиденциальности передачи данных подзадачи. В работе же предлагается подход, основанный на локальном дифференциальном методе «Deep Reinforcement Learning» в IoT на базе облачных вычислительных сервисов для обеспечения не только безопасности данных, но и индивидуальной защиты конфиденциальности путем создания разных уровней шума. Проведенные эксперименты для оценки эффективности мер безопасности, влияния фактора дисконтирования на завершение задач, экономической эффективности и возможностей защиты конфиденциальности показывают положительные результаты.

1.10.2.4 Облачная обработка больших данных на основе базы распределенного реестра

Облачная обработка больших данных на основе базы распределенного реестра (далее ТРР) – это процесс обработки, хранения и управления большими объёмами данных в распределенной среде на основе технологии цифрового реестра. Данные, хранящиеся в ТРР, распределяются по всей сети и копируются на различных узлах, в то время как система контролирует и подтверждает каждую операцию. Это позволяет улучшить безопасность, прозрачность и эффективность обработки данных, особенно в случаях, когда необходима высокая степень защиты данных или, когда требуется аудит и отслеживаемость изменений. ТРР используется в различных областях, включая финансовые услуги, логистику, здравоохранение, право и другие.

Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших данных на основе базы распределенного реестра» за последние 10 лет показана на рисунке 1.10.7 [1, 2].



Рисунок 1.10.7 – Научно–исследовательская и публикационная активность, связанная с направлением «Облачная обработка больших данных на основе базы распределенного реестра» за последние 10 лет

В работе [229] рассматривается технология блокчейн и ее возможности в различных областях, таких как Интернет вещей, облачные вычисления и социальные сети. Блокчейн предоставляет возможности для безопасного и прозрачного обмена данными, а также решения проблем, связанных с безопасностью, доверием и масштабируемостью в распределенных системах. В статье отмечаются вызовы и проблемы, связанные с применением блокчейна в различных сферах, а также представлены примеры исследований, проведенных в каждой области.

В работе [230] утверждается блокчейн и большие данные являются перспективными технологиями и занимают доминирующие позиции в списке приоритетов у многих фирм. Их внедрение ожидается значительно изменить способы ведения бизнеса и работы фирм. Блокчейн при этом предоставляет метод согласования, позволяющий создавать след аудита всех транзакций, а также обеспечивает исполнение договоренностей между сторонами сделок. Большие данные считаются безопасными и надежными в рамках сетевой архитектуры, а использование блокчейна добавляет дополнительный слой данных для процесса анализа больших данных. Все эти факторы делают блокчейн и большие данные важными направлениями в технологической сфере, которые подлежат дальнейшему развитию и исследованию.

В работе [231] рассматривает использование IoT и Smart Cities как парадигмы вычислительной технологии, которая позволяет получать информацию о состоянии города практически в реальном времени, включая уровни загрязнения. Fog Computing дополняет эту парадигму, позволяя вычисления выполняться на месте, где производятся данные сенсоров, обеспечивая низкую задержку и толерантность к отказам. Одним из вызовов при таком

подходе является гарантирование пригодности входящих данных, так как они могут быть ошибочными. Здесь на помощь приходит машинное обучение, которое может исправлять данные и производить прогнозы будущих значений. В статье представлены три вклада, направленные на решение этой проблемы, включая прогнозирование в распределенной среде с использованием набора данных дорожного движения, оценку выбросов от судов и создание новых переменных на основе данных трассировки судов, полученных из AIS. В целом, представленная в статье методика позволяет распределять и синхронизировать процессы обучения и обеспечивает устойчивость к отказам в сети при использовании устройств с низким энергопотреблением.

1.10.3 Системы обработки больших объёмов данных с использованием машинного обучения

Системы обработки больших объёмов данных с использованием машинного обучения – это комплекс алгоритмов, технологий и программных решений, предназначенный для анализа, обработки и извлечения полезной информации из массивов данных, размер которых существенно превышает возможности классических инструментов и методов. Эти системы используют принципы машинного обучения для автоматического выявления закономерностей, классификации, кластеризации и прогнозирования.

Основными тенденциями развития систем обработки больших объёмов данных с использованием машинного обучения являются:

- инструменты машинного обучения для интеграции и управления большими данными,
- расширенные инструменты аналитики и визуализации,
- распределенные вычисления и обработка данных.

1.10.3.1 Инструменты машинного обучения для интеграции и управления большими данными

Инструменты машинного обучения играют важную роль в интеграции и управлении большими данными. Они помогают оптимизировать и автоматизировать процессы обработки данных, обнаруживать скрытые закономерности и связи между данными, а также прогнозировать будущие события и тренды на основе имеющейся информации. Алгоритмы машинного обучения полезны для сбора, анализа и интеграции данных в крупных организациях. Они могут быть реализованы во всех элементах работы с большими данными, включая маркировку и сегментацию данных, анализ данных и моделирование сценариев.

Научно–исследовательская и публикационная активность, связанная с направлением «Инструменты машинного обучения для интеграции и управления большими данными» за

последние 10 лет показана на рисунке 1.10.7 [1, 2].



Рисунок 1.10.7 – Научно–исследовательская и публикационная активность, связанная с направлением «Инструменты машинного обучения для интеграции и управления большими данными» за последние 10 лет

В работе [232] разработана семантическая интеграция разнородных источников данных с использованием моделирования доменных знаний на основе онтологий для раннего обнаружения COVID–19. Существуют различные подходы к интеграции данных, такие как онтологии, методы машинного обучения, глубокое обучение и нечеткая логика. В данном тексте предложен метод автоматической онтологической интеграции данных, который может быть применен в здравоохранении. Модель состоит из: автоматического сопоставления данных, генерации локальных онтологий, комбинировании локальных онтологий и запросов к различным базам данных для получения семантически аналогичных записей. На основе данных о медицинских записях, результатах рентгеновских снимков груди и опросника симптомов COVID–19 больных из трех источников данных (SQL, mongodb и excel) были выделены пациенты, у которых есть умеренный/высокий риск развития тяжелой формы COVID–19 с достаточной точностью.

В работе [233] описаны проблемы интеграции систем обработки данных и машинного обучения в прецизионной медицине. Основная цель прецизионной медицины состоит в объединении и интеграции обширных корпусов различных баз данных о молекулярном и экологическом происхождении болезней в аналитические рамки, что позволяет разрабатывать индивидуализированные контекстно–зависимые диагностические и терапевтические подходы. Искусственный интеллект и машинное обучение используются для создания аналитических моделей сложных заболеваний, направленных на прогнозирование персонализированных условий здоровья и исходов. Эффективная интеграция и управление данными являются ключевыми компонентами успешного применения интеллектуальных

аналитических методов в медицине. Однако, при разработке таких методов имеются сложности, связанные с производительностью, вычислительной мощностью, биоэтическими ограничениями, а также ограничениями времени.

В работе [234] проведен обзор системы поддержки принятия решений по интеллектуальному планированию орошения с использованием подходов машинного обучения. В работе пришли к выводу, что использование аналитики больших данных и интеллектуальных методов обучения может существенно улучшить управление орошением для устойчивого сельского хозяйства. Решения, основанные на аналитическом моделировании, могут помочь точнее определить потребность в воде для различных культур, а системы поддержки принятия решений на основе больших данных могут обеспечить более эффективное управление капельным орошением и расчета нужной дозы воды.

1.10.3.2 Расширенные инструменты аналитики и визуализации

Использование машинного обучения в инструментах аналитики и визуализации больших данных помогает компаниям и организациям получать более точные и релевантные данные, что в свою очередь помогает им принимать более эффективные и обоснованные решения.

Научно–исследовательская и публикационная активность, связанная с направлением «Расширенные инструменты аналитики и визуализации» за последние 10 лет показана на рисунке 1.10.8 [1, 2].



Рисунок 1.10.8 – Научно–исследовательская и публикационная активность, связанная с направлением «Расширенные инструменты аналитики и визуализации» за последние 10 лет

В работе [235] рассмотрено проектирование и реализация системы реконструкции многомерной визуализации старых городских пространств на основе нейронных сетей. В данной статье представлено исследование и анализ конструкции модели свёрточной

нейронной сети и системы многомерной визуализации старых городских пространств, а также предложен дизайн многогранной системы реконструкции визуализации старых городских пространств на основе нейронной сети. В процессе обучения переконфигурируемой сверточной нейронной сети с использованием наборов данных используется неглубокая структура сети, чтобы уменьшить количество параметров, обучить модель более общим особенностям. Если количество наборов данных невелико, то используется способ увеличения размера наборов данных и повышения точности распознавания переконфигурируемой сверточной нейронной сетью. Для сценариев больших данных был предложен и реализован метод обновления многогранной визуализации данных в режиме реального времени, который позволяет снизить загрузку и задержку сети, вызванные изменениями графиков многомерных данных, снизить ошибки данных и поддерживать стабильность системы в сценарии конкуренции в старом городском пространстве.

В работе [236] разработана система анализа пространства визуальных параметров с помощью глубокого обучения для космологических симуляций. Космологи часто создают математические модели симуляции, чтобы изучать наблюдаемую вселенную. Однако запуск высококачественной симуляции занимает много времени и может затруднить анализ. Это особенно актуально в тех случаях, когда анализ включает в себя проверку большого количества конфигураций входных параметров симуляции. Поэтому выбор нужной конфигурации входных параметров стал важной частью процесса анализа. В данной работе мы предлагаем интерактивную визуальную систему, которая эффективно помогает пользователям понять пространство параметров, связанных с их космологическими данными. Разработанная система использует модель на основе GAN, чтобы перестроить выводы симуляции без запуска дорогостоящей симуляции. В работе также извлекается информация, полученная из глубоких нейронных сетей на основе заменительной модели, чтобы упростить исследование пространства параметров. В результате этих исследований были получены ценные настройки входных параметров симуляции и анализированы подрегионы.

В работе [237] проведена оценка производительности управления данными интеллектуальных алгоритмов для экосистемы больших данных. В работе оценили и сравнили несколько алгоритмов машинного обучения для анализа качества платформы, сравнили Apache Spark ML-lib с Rapid Miner и Sklearn, которые являются двумя дополнительными платформами обработки больших данных и машинного обучения. Логистический классификатор, классификатор «дерево решений», классификатор «случайный лес» и градиентный бустинг классификатор – это четыре алгоритма машинного обучения, которые сравниваются между платформами. Кроме того, в работе были проведены тестирование общих методов регрессии, таких как линейный регрессор, дерево решений регрессор,

случайный лес регрессор и градиентный бустинг регрессор на наборах данных «SUSY» и «Higgs». Также были оценены методы неконтролируемого обучения, такие как K-средних и модели смеси Гаусса на наборе данных SUSY и Hepmass, чтобы определить устойчивость PySpark по сравнению с моделями классификации и регрессии. Использовались наборы данных «SUSY», «HIGGS», «BANK» и «HEPMASS» из репозитория данных UCI.

1.10.3.3 Распределенные вычисления и обработка данных

Распределенные вычисления являются базовой технологией для кластерных вычислений и облачных вычислений. Это позволяет вычислять большие наборы данных, разделяя их на мелкие части по узлам. Эти узлы связаны через коммуникационную сеть и работают как единая вычислительная среда для решения конкретной задачи.

Научно–исследовательская и публикационная активность, связанная с направлением «Распределенные вычисления и обработка данных» за последние 10 лет показана на рисунке 1.10.9 [1, 2].



Рисунок 1.10.9 – Научно–исследовательская и публикационная активность, связанная с направлением «Распределенные вычисления и обработка данных» за последние 10 лет

В работе [238] рассмотрено эффективное обнаружение выбросов с помощью модели прогнозирования финансового кризиса на основе глубокого обучения в среде больших данных. В данном исследовании разработан новый метод прогнозирования финансовых кризисов на основе комбинации технологий анализа больших данных, Интернета вещей, облачных вычислений и машинного обучения. Для эффективного прогнозирования используется новый алгоритм выбора оптимального набора признаков, который помогает улучшить результаты классификации. Экспериментальная проверка алгоритма показала его превосходство над существующими методами. В целом, данное исследование поможет бизнесам и странам предотвратить финансовые кризисы и сильнее утвердить свою

экономическую позицию.

В работе [239] рассмотрены методы аналитики больших данных в управлении цепочками поставок с использованием машинного обучения. Использование аналитики больших данных позволяет эффективно управлять бизнес–процессами в различных сферах, включая бухгалтерию, маркетинг, логистику и операции. С развитием машинного обучения и вычислительной инфраструктуры важность аналитики больших данных для управления логистическими процессами возрастает. Как показывает статистика, на 2022 год многие компании не проявляются интерес к данной технологии из–за сложности разработки и нехватки специалистов.

В работе [240] рассмотрены статистические методы и методы машинного обучения для анализа транскриптомных данных с пространственным разрешением. Недавние прорывы в технологии пространственной транскриптомики позволяют многократный профилинг транскриптомов клеток и их пространственное расположение. С улучшением ёмкости и эффективности опытных технологий возникает необходимость развития аналитических подходов. Одновременно с непрерывной эволюцией протоколов секвенирования требуется повторно оценить предположения существующих аналитических методов и адаптировать их к растущей сложности данных. В работе продемонстрирован последний прогресс в статистических и машинных методах пространственной транскриптомики, а также подводятся итоги полезных ресурсов и обозначаем сложности и возможности, которые стоят перед исследователями.

1.10.4 Вывод

Основными тенденциями развития систем обработки больших объёмов данных являются инструменты машинного обучения для интеграции и управления большими данными и распределенные и облачные вычисления в режиме реального времени.

2 АНАЛИЗ СУЩЕСТВУЮЩИХ И ПЕРСПЕКТИВНЫХ РЫНКОВ, РАЗВИВАЮЩИХСЯ С ИСПОЛЬЗОВАНИЕМ РЕЗУЛЬТАТОВ НАУЧНОЙ И (ИЛИ) НАУЧНО–ТЕХНИЧЕСКОЙ ДЕЯТЕЛЬНОСТИ В РОССИЙСКОЙ ФЕДЕРАЦИИ И ЗА РУБЕЖОМ

2.1 Искусственный интеллект в компьютерном зрении

Российский рынок искусственного интеллекта в компьютерном зрении активно развивается. Данный рост вызван такими факторами, как более быстрая обработка и более высокая точность в сочетании с экономическими преимуществами системы компьютерного зрения [241]. Однако высокая потребность в высококвалифицированных специалистах и заботах о конфиденциальности и безопасности пользовательских данных в первую очередь сдерживает рост рынка. Более того, растущее применение компьютерного зрения в непромышленных приложениях и достижения в области компьютерного зрения и искусственного интеллекта оказывают положительное влияние на тенденции рынка компьютерного зрения.

В 2022 году на мировом рынке компьютерного зрения доминировал промышленный сегмент рынка, и ожидается, что он сохранит свое доминирование в предстоящие годы. Промышленный сегмент включает вертикали, которые включают приложения компьютерного зрения в производственных процессах, таких как автомобилестроение, бытовая электроника, металлы, деревообработка, бумаги, пищевая промышленность, в том числе пищевая упаковка, машиностроение и текстиль и др.

Глобальная индустрия компьютерного зрения в основном классифицируется на основе компонентов, продуктов, приложений, вертикали и региона. В зависимости от компонентов рынок подразделяется на аппаратное обеспечение, программное обеспечение и услуги. Кроме того, аппаратный сегмент подразделяется на камеры, устройства захвата кадров, оптику, светодиодное освещение и процессоры, в то время как программное обеспечение подразделяется на традиционное компьютерное зрение и глубокое обучение. В зависимости от продукта рынок подразделяется на системы компьютерного зрения на базе ПК и интеллектуальных камер, в то время как в зависимости от приложения рынок подразделяется на обеспечение качества и инспекцию, позиционирование и руководство, измерение, идентификацию, профилактическое обслуживание. По вертикали рынок разделяется на промышленный и непромышленный.

Как видно из рисунка 2.1.1, COVID–19 в 2019–2022 годах оказал смешанное влияние на рынок компьютерного зрения. С одной стороны, во многих сферах промышленности был экономический спад в производстве и цепочке поставок, что негативно сказалось на

возможности внедрения новых технологий на основе компьютерного зрения. С другой стороны, появилось в данный период множество непромышленных применений данной технологии, например, мониторинг состояния дорог, рентгеновский анализ, диагностика новообразований и патологий, измерение кровопотери, анализ движения, мониторинг урожая, розничная торговля, автоматическое оценка наполненности складов, интеллектуальная видеоаналитика, подсчет людей, выявление у людей аномальных признаков и др. [242]. После COVID–19 прогнозируемый объём мирового рынка компьютерного зрения в 2022 году составил 10 813 миллионов долларов, а к 2030 году, по прогнозам, достигнет 41 117 миллионов долларов [243].

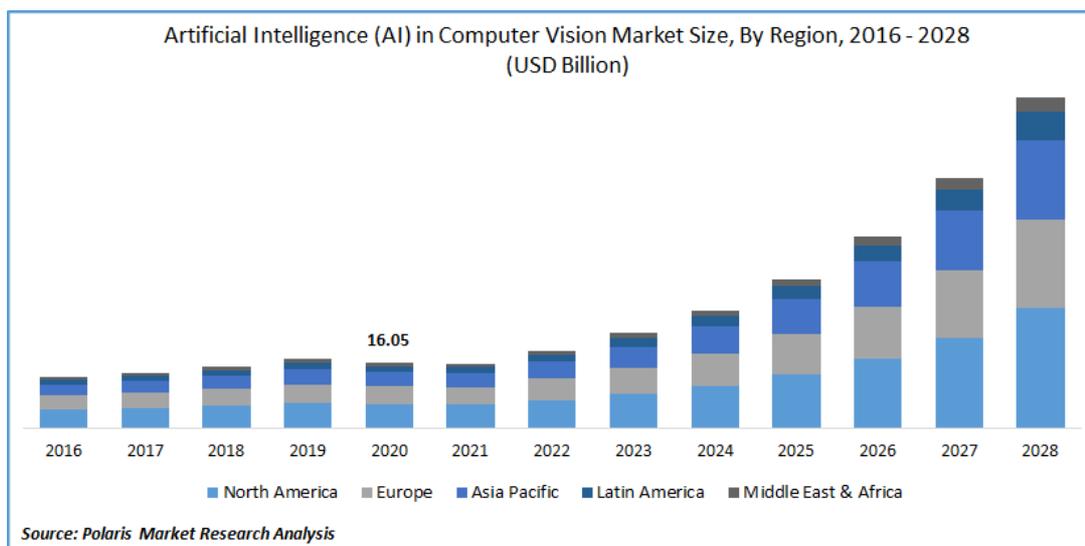


Рисунок 2.1.1 – Прогнозируемый рост рынка компьютерного зрения до 2030 года [244]

В настоящее время можно выделить 9 компаний, занимающих больше половины рынка компьютерного зрения в мире:

— Корпорация «Cognex» – американский производитель систем технического зрения, программного обеспечения и датчиков, используемых в автоматизированном производстве для проверки и идентификации деталей, обнаружения дефектов, проверки сборки изделия и руководства сборочными роботами;

— Корпорация Google – американская многонациональная технологическая компания, специализирующаяся на онлайн–рекламе, технологии поисковых систем, облачных вычислениях, компьютерном программном обеспечении, квантовых вычислениях, электронной коммерции, искусственном интеллекте и бытовой электронике. Продуктом в области компьютерного зрения компании Google является полностью управляемая среда разработки «Vision AI» для создания собственных приложений компьютерного зрения или получения информации из изображений и видео с помощью предварительно подготовленных API, AutoML или пользовательских моделей;

— Корпорация Microsoft – американская транснациональная технологическая

корпорация, занимающаяся разработкой программной продукции и аппаратных продуктов. Microsoft была разработана система компьютерного зрения, способная распознавать широкий спектр объектов, текста, лиц, машин, анализировать видео в режиме реального времени и многое другое. При этом компания снизила порог вхождения в свой продукт на столько, что потребителю практически не требуется опыт в области машинного обучения для создания собственной системы компьютерного зрения;

— Корпорация «Intel» – американская транснациональная корпорация, один из крупнейших в мире производителей полупроводниковых чипов. Крупнейший поставщик компьютерного зрения для приложений промышленности;

— Корпорация «KEYENCE» – японская компания, занимающаяся организацией прямых продаж. Компания разрабатывает и производит датчики автоматизации, системы визуального наблюдения, считыватели штрих-кодов, лазерные маркеры, измерительные приборы и цифровые микроскопы;

— Корпорация «Matterport» – американская компания, предоставляющая платформу для захвата, хранения, обмена 3D-изображениями и моделями. основными услугами компании являются камеры Matterport и облако Matterport;

— Корпорация «NATIONAL INSTRUMENTS» – американская многонациональная компания, являющаяся производителем автоматизированного испытательного оборудования и программного обеспечения для виртуальных приборов. Распространенные приложения включают в себя сбор данных, управление приборами и техническое зрение;

— Корпорация «Omron» – японская корпорация, производитель электроники и один из мировых лидеров в производстве средств автоматизации. Одним из успешных продуктов компании является системы ночного видения Omron, объединяющая камеры с техническим зрением и программное обеспечение для обработки изображений;

— Корпорация «Sony Semiconductor Solutions Corporation» – японская корпорация, является одним из крупнейших в мире производителей бытовой и профессиональной электроники, крупнейшим производителем игровых приставок и крупнейшим издателем видеоигр. Компания Sony, занимающая 55 % долю на рынке датчиков изображения. Компания активно работает над технологиями 3D компьютерного зрения, такими как оценка глубины, визуальный щелчок и алгоритмы 3D-моделирования в камерах.

Объём рынка искусственного интеллекта вместе с компьютерным зрением в России по итогам 2022 г. составил около 635 млрд руб., что на 15 % больше, чем годом ранее. При этом отрасль розничной торговли является одним из лидеров по уровню проникновения ИИ [245]. Эти технологии применяются для автоматизации бизнес-процессов и увеличения продаж за счет персонализации сервисов. По прогнозам аналитиков, в 2023 году российский рынок

компьютерного зрения может достигнуть 38 млрд руб., что практически в 5 раз больше, чем в 2019 году [246].

Согласно аналитическому отчету Центра компетенций НТИ на базе МФТИ по направлению «Искусственный интеллект», работало более 400 компаний, из них 75 – в области компьютерного зрения. Большую долю объёма российского рынка занимают следующие компании:

- «Яндекс,
- АBBY,
- ПАО «Сбербанк»,
- VisionLabs,
- Группа ЦРТ
- VK.

Яндекс – транснациональная компания, занимающаяся множеством приложений компьютерного зрения. Например, разработкой систем управления беспилотным автомобилем, которая использует лидары, камеры, радары, GPS и IMU (гиростабилизатор) для достижения пятого уровня автономности (полной автономности автомобиля). Также существует сервис «Yandex Vision», который занимается решением бизнес-задач, связанных с распознаванием текста, документов, автомобильных номеров, обнаружением лиц и классификации изображений.

VisionLabs – мировой лидер в области распознавания лиц. VisionLabs разрабатывает технологии и продукты для финансовой сферы, ритейла, видеонаблюдения и безопасности. Компания специализируется на создании программных решений и сервисов на базе технологий компьютерного зрения. Основными продуктами компании являются:

- «VisionLabs Luna» – сервис для видеоаналитики, использующий инновационную технологию распознавания образов. Данная технология входит в тройку лучших в мире по качеству распознавания в реальных условиях. Это подтверждено тестами на независимой базе данных University of Massachusetts – Labeled Faces in the Wild;

- «VisionLabs Thermo» – сервис на основе компьютерного зрения, способный бесконтактно обеспечить эпидемиологический контроль и автоматически выявляют людей с повышенной температурой в режиме реального времени;

- «LUNA Cars» – данный продукт способен определить большинство атрибутов транспортного средства: государственный номер, марку, модель, цвет, тип и категорию;

- «VisionLabs Luna Pass» данный сервис позволяет проводить фотографирование через веб-приложение. При фотографировании клиентов для обслуживания или оформления заявок система проверяет качество полученной фотографии по более чем 15 параметрам,

среди которых отсутствие засветов, достаточная освещенность, поворот головы и т.д., а также контролирует, чтобы процесс проходил без подмены изображения.

ABBYY – международная компания, разработчик решений в области интеллектуальная обработка документов, системы потокового ввода документов и данных, анализ бизнес–процессов и технологии оптического распознавания символов. В 2019 году ABBYY и МФТИ открыли совместную научную лабораторию компьютерного зрения и обработки естественного языка. Основными продуктами компании в области компьютерного зрения являются:

— «ABBYY FlexiCapture» сервис для извлечения данных из неструктурированных документов, таких как договоры, уставы, доверенности, письма или тексты новостей. Платформа умеет находить в текстах документов важные сущности, такие как имена, названия организаций и адреса, а также выявлять значимые факты и связи между ними. С дополнительными функциональными модулями NLP ABBYY FlexiCapture юристы смогут в разы быстрее анализировать многостраничные договоры, выявлять значимые условия сделки и создавать карточки документов в корпоративной системе;

— «ABBYY FineReader Server» – корпоративное серверное решение для распознавания документов и преобразования файлов в PDF и другие форматы. Решение быстро и точно распознает текстовые изображения и переводит их в PDF, PDF/A, Microsoft Word или другие типы файлов, удобные для поиска, просмотра, обмена и хранения;

— «ABBYY PassportReader» – совместный продукт с VisionLabs для распознавания образов в реальном времени идентифицирует личность человека по фото или кадру из видеопотока. Оно мгновенно сопоставляет образ клиента с его данными из информационной системы организации, а также подтверждает подлинность предоставленных им документов.

ПАО «Сбербанк» – российский финансовый конгломерат, крупнейший универсальный банк России и Восточной Европы, зарегистрированный под торговой маркой «Сбер». Сбер AI – единая площадка, где собрано всё об искусственном интеллекте с упором на технологии, продукты и экспертов Сбера. Лаборатория искусственного интеллекта – стратегическое подразделение и главный центр компетенций Сбера в области искусственного интеллекта, в том числе компьютерного зрения. Основными продуктами компании в области компьютерного зрения являются:

— «SaluteVision Mobile SDK» – набор инструментов, позволяющий российскому бизнесу и государственным организациям в короткие сроки встроить системы компьютерного зрения в свои мобильные приложения на Android и iOS [247];

— «Smartix» – совместный проект Сбера и VisionLabs для оплаты взглядом на кассах самообслуживания. Система распознавания лиц разработана на базе технологий

компании VisionLabs. С помощью камер система компьютерного зрения обнаруживает лицо человека, выбирает из всего потока самый удачный кадр, извлекает из него биометрический шаблон и сравнивает с внутренней базой. Если совпадение находится, то транзакция одобряется, и оплата проходит;

— «SberShuttle» – роботизированную систему автоматизации складской логистики. Обновленное ПО позволяет согласованно управлять всем парком роботов. Благодаря переходу на собственные программные продукты по управлению SberShuttle в 2022 году надёжность работы системы повысилась на 30 %, сократилось время на перемещение товаров до станции сборки заказов, а также оптимизированы потоки товаров и распределение задач с точки зрения приоритизации, что позволило сократить среднее время на сборку одного заказа на 10–15 % [248].

Группа ЦРТ – российская компания, специализирующаяся на инновационных системах в сфере распознавания лиц, голосовой биометрии, распознавания и синтеза речи, многоканальной записи, обработки и анализа, аудио– и видеоинформации. Одним из успешных сервисов компании является «Визирь». «Визирь» – система видеонаблюдения и автоматической идентификации личности по биометрическим характеристикам лица. Системы биометрического контроля доступа, учёта и поиска «Визирь» используется в местах массового скопления людей для обнаружения лиц в реальном масштабе времени, ведения структурированной картотеки фотоизображений и сопутствующей информации, поиска в картотеке по фотоизображению и сопутствующей информации, а также в видеоархиве.

VK – российская инвестиционная технологическая корпорация, основанная в 1998 году. VK на базе своей универсальной облачной платформы для развития цифровых сервисов «VK Cloud» разработал систему компьютерного зрения «Mail.ru Vision» для распознавания лиц и объектов на базе машинного обучения.

Группа ЦРТ – российская компания, специализирующаяся на инновационных системах в сфере распознавания лиц, голосовой биометрии, распознавания и синтеза речи, многоканальной записи, обработки и анализа, аудио– и видеоинформации.

Потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части искусственного интеллекта в компьютерном зрении являются крупные и средние компании занимающиеся обрабатывающей, автомобильной, пищевой, химической и нефтеперерабатывающей, лесной промышленностью, органы здравоохранения, медицинские учреждения, занимающейся диагностикой онкологий и новообразований, розничной торговлей, такие как: ПАО «КАМАЗ», АО «Газпромнефть», X5 Group, ПАО «Магнит», ООО «Вайлдберриз», холдинг Ozon, НМИЦ онкологии им. Н.Н. Блохина и др.

Также стоит выделить инновационный центр. «Сколково» - действующий в Москве современный научно-технологический инновационный комплекс по разработке и коммерциализации новых технологий, первый в Российской Федерации строящийся с нуля наукоград. Из заметных проектов стоит выделить платформу Retailix.Ai. Retailix.Ai – это платформа, которая предлагает цифровые решения для автоматизации операций в торговых сетях. Сервисы на этой платформе были применены в крупных розничных сетях и показали высокую эффективность, а теперь они доступны для использования малыми и средними торговыми сетями с точки зрения стоимости и простоты внедрения. Применение некоторых из цифровых решений обеспечивает рост товарооборота в крупных сетях от 4 % до 6 %. Также исследовательский центр проводит исследования в сфере искусственного интеллекта по направлению оптимизации управленческих решений в целях снижения углеродного следа и создание моделей и фреймворков ИИ для решения задач устойчивого развития промышленности и экономики России.

Университет Иннополис — российская автономная некоммерческая организация высшего образования в городе Иннополисе (Верхнеуслонский район, Республика Татарстан), специализирующаяся на образовании, исследованиях и разработках в области информационных технологий и робототехники. Среди всех проектов можно выделить наиболее крупные:

- Программно-аппаратный комплекс контроля износа режущего инструмента на базе технологий искусственного интеллекта (AI) и компьютерного зрения (CV),
- Разработка опытного образца системы помощи водителю автобуса в части детекции объектов и их трекинга,
- Разработка программного комплекса группового управления мобильными коллаборационными платформами для выполнения логистических и вспомогательных функций в замкнутых пространствах существенной протяженности,
- Разработка алгоритмов управления и обработки сенсорной информации для автономного пилотирования транспортных средств в условиях бездорожья,
- Система автоматического распознавания патологий легких по рентгеновским изображениям с применением технологий ИИ "AI RADIOLOGY",
- Исследование качества распознавания поголовья коров с помощью машинного зрения.

ФГАОУ ВО "Национальный исследовательский университет – Высшая школа экономики" занимается межотраслевыми технологиями искусственного интеллекта и искусственный интеллект для иных приоритетных отраслей экономики и социальной сферы. Основными проектами Высшей школы экономики являются:

- Интеллектуальная автоматизация ручных операций, распознавание операций на производстве и интеллектуальные методы в промышленной безопасности,
- Экспериментальные методы оценки качества продукции систем ИИ,
- Диагностические и ассистивные речевые технологии на основе ИИ.

Центр компетенций НТИ МФТИ - это платформа, созданная в рамках национальной технологической инициативы (НТИ) и функционирующая на базе Московского физико-технического института (МФТИ). Его задача заключается в развитии инновационных проектов и технологий в области искусственного интеллекта, квантовых технологий, нанотехнологий, биотехнологий, космических технологий и других областей. Центр также обучает специалистов в этих областях и участвует в международных научных и технологически проектах. Основными проектами центра компетенций НТИ являются:

- Новая интерактивная биометрия: разработка принципиально новых технологий биометрической идентификации пользователей с помощью анализа и интеллектуальной обработки рефлекторных реакций человека,
- Технологии распознавания околоземных космических объектов для роботизированных телескопов: создание сети оптических телескопов-роботов, оптических матриц и алгоритмов для обнаружения, сопровождения, распознавания и классификации околоземных космических объектов,
- Системы технического зрения для беспилотных автомобилей: разработка системы технического зрения для интеллектуальной обработки фото и видеoinформации с использованием сверточных нейросетей, глубокого обучения и программно-аппаратной интеллектуальной системы поиска, локализации и классификации объектов на фото и видео в режиме реального времени.

Как показывает статистика, рынок технического зрения будет расти и расширяться как в мире, так и в России. В работе были выявлены основные ключевые российские и зарубежные разработчики технического зрения. В настоящее время в России существуют готовые отечественные решения, которые могут решить широкий спектр задач начиная от промышленности заканчивая розничной торговлей. Был оценён объём российского и зарубежного рынка, а также определены перспективы роста более чем на 5 лет вперед.

2.2 Искусственный интеллект в обработке естественного языка

Российский рынок искусственного интеллекта в обработке естественного языка находится в стадии активного развития. Данный рост вызван повышенным спросом в области информационных технологий и телекоммуникаций, банковского дела, розничной, электронной торговли, а также в таких сферах как: промышленность, образование,

здравоохранения и др.

Объём мирового рынка обработки естественного языка (NLP) оценивался в 19,68 млрд долларов США в 2022 году и, по прогнозам, вырастет с 24,10 млрд долларов США в 2023 году до 112,28 млрд долларов США к 2030 году, демонстрируя среднегодовой рост в 24,6 % в течение прогнозируемого периода [249].

Пандемия COVID–19 негативно повлияла на мировую экономику, поскольку правительства по всему миру ввели карантин и ограничения. Однако пандемия COVID–19 подтолкнула предприятия по всему миру к переходу на цифровые платформы и внедрению решений на основе цифровых технологий для оптимизации своих бизнес–операций. Предприятия внедрили цифровые стратегии для автоматизации своих бизнес–процессов. В глобальном масштабе объём рынка незначительно увеличился на 0,2 % [249].

По оценкам, в 2022 году лидирующая доля локального сегмента составит 59,8 % [250]. Локальное развертывание распознавания естественного языка обеспечивает полный контроль, видимость и безопасность аутентификации данных. Кроме того, его легче масштабировать в соответствии с корпоративным спросом и повышать эффективность благодаря встроенному резервированию.

Способность облачных решений обрабатывать огромные массивы данных и обеспечивать улучшенный пользовательский опыт побудила ряд компаний сделать выбор в пользу облачного развертывания, а не локального. Поскольку большинство предприятий не имеют инфраструктуры и сетей, способных обрабатывать большие наборы данных, на рынке NLP существует огромный спрос на развертывание в облаке.

В 2022 году заметен повышенный спрос на облачную обработку естественного языка во всем мире. Ожидается, что расширение внедрения облачного NLP будет способствовать расширению рынка. Ожидается, что в облачном сегменте будет наблюдаться самый быстрый рост за прогнозируемый период.

Одним из ключевых сдерживающих факторов развития рынка NLP является проблемы безопасности данных и затрат, связанных с кибербезопасностью.

Ключевыми игроками на мировом рынке NLP являются следующие компании:

— Google – американская многонациональная технологическая компания, которая внедрила NLP во всю свою экосистему. Также компания владеет облачными сервисами для обработки и генерации естественного языка «Cloud Natural Language» и «Bard» соответственно;

— Amazon Web Services, Inc. – американская многонациональная технологическая компания, владеет сервисом «Amazon Comprehend» для обработки естественного языка;

— Корпорация Microsoft – американская транснациональная технологическая

корпорация, владеет облачным сервисом для обработки естественного языка «Azure Cognitive Service for Language»;

— Корпорация International Business MachinesI (далее IBM) – американская многонациональная технологическая корпорация, владелец сервиса для обработки естественного языка «IBM Watson»;

— OpenAI – американская исследовательская лаборатория искусственного интеллекта (ИИ), состоящая из некоммерческой «OpenAI Incorporated» и ее коммерческой дочерней компании «OpenAI Limited Partnership». Основными продуктами компании в области генерации и обработки естественного языка являются сервисы «OpenAI Codex», «GPT-4» и «GPT-3»;

— SoundHound AI, Inc. – американская компания, занимающаяся распознаванием, обработкой естественного языка. Владеет платформой голосового искусственного интеллекта «SoundHound»;

— SAS Institute Inc. – американский многонациональный разработчик программного обеспечения для аналитики. Компания предоставляет такие услуги, как: категоризация контента, классификация, корпусной анализ, контекстное извлечение, анализ настроений, преобразование текст в речь и наоборот, краткое изложение документа, машинный перевод;

— Automated Insights – американская технологическая компания, специализирующаяся на программном обеспечении для генерации естественного языка, которое превращает большие данные в удобочитаемый, структурированный текст. Продуктом компании является сервис для генерации естественного языка «Wordsmith».

Объём российского рынка обработки естественного языка на 2022 год примерно оценивается в 142 млн. долларов США, также по оценке аналитиков рост рынка составляет около 38 % [251].

Крупнейшими компаниями в Российской Федерации, занимающиеся разработкой и инвестициями в обработку естественного языка являются: Яндекс, Группа ЦРТ, АВВУУ, VK, Just AI, PROMT, Тинькофф, Наносемантика, BrandAnalytics, Сбербанк, RCO, АСМ Решения, Медиалогия, Kribum, МТС, Naumen. Рассмотрим наиболее крупных представителей рынка.

Яндекс – международная компания в отрасли информационных технологий, ведущая поисковая система в интернете в России по количеству обрабатываемых поисковых запросов. Главными продуктами в области генерации и обработки естественного языка являются: «Яндекс Переводчик», «Яндекс Алиса», «Яндекс SpeechKit».

В «Яндекс переводчике» для NLP работает одновременно нейронная сеть, статистическая модель, и специальный алгоритм выбора перевода, что обеспечивает точный

и понятный пользователю перевод. На 2023 год «Яндекс Переводчик» поддерживает 90 языков, включая специфичные языки: эльфийский, эсперанто, эмодзи.

«Яндекс Алиса» используется NLP и машинное обучение для обработки голосовых и текстовых запросов, для ответов на запросы пользователей «Яндекс Алиса» использует технологии генеративных нейронных сетей, таких как LSTM (Long Short–Term Memory) и GPT (Generative Pre–trained Transformer). С 2023 года «Яндекс Алиса» для генерации естественного языка также использует технологию «GPT–4».

«Яндекс SpeechKit» – это сервис распознавания речи, разработанный компанией Яндекс. Он позволяет преобразовывать аудиозаписи в текст и наоборот, а также синтезировать голосовую речь на разных языках. Для распознавания речи данный сервис использует совместно NLP и нейронную сеть, обученную на большом количестве записей речи на различных языках и диалектах.

Группа ЦРТ – российская компания, входящая в экосистему «СБЕР», специализирующаяся на разработке и производстве компьютерной техники и программного обеспечения. Наиболее успешные продукты в области обработки естественного языка: «Voice2Med» и «Speech Analytics Lab». Также компания занимается разработкой голосовых ассистентов и уже успешно разработала: для службы «122», банка ВТБ, для медицинских служб, ОАО «РЖД» и др.

ABBY – международная компания, предоставляющая широкий спектр продуктов, услуг и сервисов в области обработки естественного языка:

- «ABBY TextGrabber» – мобильное приложение для сканирования и распознавания на изображениях;
- «ABBY FineReader» – программное обеспечение для распознавания текста и конвертации документов в различные форматы;
- «ABBY FlexiCapture» – программное обеспечение для автоматического извлечения данных из различных документов;
- «ABBY Timeline» – инструмент для анализа текстов и выделения ключевых событий и фактов;
- «ABBY Compreno» – платформа для создания и обучения искусственного интеллекта с возможностью обработки естественного языка;
- «ABBY SmartCAT» – облачная платформа для переводов и управления проектами;
- «ABBY Mobile OCR Engine» – библиотека для распознавания текста на мобильных устройствах;
- «ABBY Language Services» – услуги по переводу и локализации контента на

различные языки.

VK – российская инвестиционная технологическая корпорация, предоставляющая широкий спектр продуктов, услуг и сервисов в области обработки естественного языка:

- «VK Chatbot Platform» – это инструмент для создания и настройки чат-ботов в VK, использующих технологии обработки естественного языка для взаимодействия с пользователями;

- VK Реклама использует NLP для определения тематики рекламных объявлений и их соответствия контексту, в котором они будут показаны. Алгоритмы NLP помогают автоматически классифицировать тексты объявлений, определять их ключевые слова;

- Прометей – искусственный интеллект для определения релевантного контента и продвижения его в социальной сети;

- VK HR Тек – использует NLP для анализа резюме и определения соответствия кандидатов требованиям вакансий;

- VK Капсула – умная колонка со встроенным голосовым помощником «Маруся»;

- VK Assistant – это голосовой помощник, который умеет обрабатывать естественный язык для того, чтобы пользователь мог задавать ему вопросы и получать ответы;

- Платформа для создания новостных видео – сервис для создания новостных и репортажных видео студийного качества.

Just AI – российская компания, специализирующаяся на технологиях искусственного интеллекта, машинного обучения и понимания естественного языка. Крупнейшим инструментом Just AI в области NLP является «САИЛА». Также можно выделить следующие сервисы связанные с NLP:

- Aimylogic – сервис для создания ботов с искусственным интеллектом;

- Aimyvoice – сервис для синтеза голоса;

- JAICP – это платформа для разработки разговорных решений любой сложности: AI чат-ботов, ботов для входящих и исходящих звонков, виртуальных ассистентов и навыков для них;

- Aimybox – это набор инструментов для голосового помощника с открытым исходным кодом и готовым пользовательским интерфейсом, который для создания своего собственного голосового помощника и встраивание его в любое приложение или устройство.

PROMT – российская компания, разработчик систем машинного перевода, также занимается исследованиями и разработками в области искусственного интеллекта. PROMT предоставляет весь спектр решений для автоматического перевода – универсальные и специализированные переводчики для Windows, Linux, iOS, Android, Mac OS, SDK переводчика для интеграции в программные комплексы и сервисы, а также системы анализа

текстов. Все решения компании работают без обращения в интернет или к сторонним сервисам и гарантируют безопасность конфиденциальной информации. Продукция компании включена в реестр отечественного ПО Минкомсвязи и удостоена многих российских и международных наград. Практически во всех продуктах, выпущенных ПРОМТ используется обработка естественного языка, рассмотрим некоторые из них:

— PROMT.One – это онлайн–сервис машинного перевода, который позволяет пользователям переводить тексты, веб–страницы и документы на различные языки;

— PROMT.Translate – это бот для мессенджера Telegram, который позволяет пользователям переводить тексты на разные языки прямо внутри приложения;

— PROMT Cloud API SaaS – это облачное API, которое позволяет разработчикам интегрировать технологии машинного перевода и обработки естественного языка в свои приложения и веб–сервисы;

— PROMT Expert Neural – это десктопная программа на основе нейронных сетей и NLP для перевода текстов любого объёма, документации всех популярных форматов с сохранением исходного форматирования и деловой переписки.

Тинькофф – российская компания, предоставляющая финансовые услуги, прежде всего известная своими банковскими и финансовыми продуктами. Тем не менее, они также инвестируют в технологии искусственного интеллекта и обработки естественного языка, чтобы улучшить свои услуги и качество обслуживания клиентов. Некоторые из их основных продуктов, использующих обработку естественного языка:

Голосовой помощник Олег – виртуальный голосовой помощник, который помогает клиентам управлять личными финансами, совершать транзакции и отвечать на вопросы, связанные с продуктами и услугами Тинькофф;

Tinkoff VoiceKit – сервис для преобразования текста в голос и наоборот;

Тинькофф Сервис речевой аналитики для бизнеса – это сервис, использующий распознавание и синтез речи вместе с обработкой естественного языка для определения настроения клиента и его реакцию на предлагаемые услуги. Данная технология решает следующие задачи: контроля качества работы операторов и менеджеров, выявления слабых мест в скриптах, оценки эффективности скриптов и работу сотрудников, понимания, как работают акции и спецпредложения, и увеличения продаж.

Сбербанк – крупнейший российский банк, который активно использует обработку естественного языка. Рассмотрим несколько популярных продуктов банка:

— СберСалют – семейство виртуальных ассистентов, созданных в Сбербанке;

— Сбербанк Онлайн – мобильное приложение для управления финансами, которое также использует технологии NLP для упрощения процесса взаимодействия с пользователем;

— Сбер: Антифрод-система – это система мониторинга и предотвращения мошеннических операций, которая в режиме реального времени проверяет каждый платеж в Сбере;

— Сбербанк система определения мошенников по IP – технология позволяет определить возможные факты мошенничества на основании связки между IP-адресом и географической позицией участников транзакции;

— SberBoom Умные колонки – линейка умных колонок Сбера. Данные колонки используют технологии распознавания и генерации речи и NLP для понимания естественного языка и генерации ответа;

— Сбер: GigaChat – мультимодальная версия нейросети от Сбера. Она умеет отвечать на вопросы пользователей, поддерживать диалог, создавать тексты и картинки на основе описаний;

— MedBench – платформу для решения ИИ-задач в медицине. На данной платформе уже доступны три задачи по автоматической обработке естественного языка: на логику и знания, вопросно-ответное моделирование и понимание смысла медицинского текста;

— Сбер RuGPT-3 – нейросеть обученная на русскоязычных текстах разных стилей: на энциклопедиях, социальных сетях, художественной и бизнес-литературе;

— Сбер ruRoBERTa – текстовая модель для обработки и понимания текста разработанная «SberDevices»;

— Сбер Аура – Сбер и Robotics совместно разработали сервис для защиты от мошенников на основе обработки естественного языка с точностью определения до 80 %.

Центр компетенций НТИ МФТИ - это платформа, созданная в рамках национальной технологической инициативы (НТИ) и функционирующая на базе Московского физико-технического института (МФТИ). Его задача заключается в развитии инновационных проектов и технологий в области искусственного интеллекта, квантовых технологий, нанотехнологий, биотехнологий, космических технологий и других областей. Центр также обучает специалистов в этих областях и участвует в международных научных и технологически проектах. Основными проектами центра компетенций НТИ являются:

— Языковые модели для AGI: разработка и исследование моделей машинного обучения для систем обработки естественного языка (DeepLanguage),

— Разговорный искусственный интеллект DeepPavlov: разработка и исследование моделей машинного обучения и использование их для создания доменных разговорных навыков, на основе размеченных данных и антологий, а также ведения диалога на общие темы – проект DeepPavlov,

— Инструменты анализа мультимодальных данных TopicNet: разработка

инструментов анализа мультимодальных текстовых и транзакционных данных - проект «TopicNet»,

— Психолингвистический анализ информационного пространства: разработка комплекса программного обеспечения, лингвистических и поведенческих моделей и методик анализа информационного пространства социальных медиа для автоматического анализа неструктурированных текстов из социальных сетей и СМИ.

Потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно-технической деятельности в Российской Федерации в части искусственного интеллекта в обработке естественного языка являются крупные и средние компании занимающиеся банковским делом, электронной торговлей, розничной торговлей, а также государственные органы, министерство финансов, здравоохранения, образования и науки и др.

Как показывает статистика, рынок обработки естественного языка будет расти и расширяться как в мире, так и в России. В работе были выявлены основные ключевые российские и зарубежные разработчики технического зрения. В настоящее время в России существуют готовые отечественные решения, которые могут решить широкий спектр задач начиная от промышленности заканчивая розничной торговлей. Был оценён объём российского и зарубежного рынка, а также определены перспективы роста более чем на 5 лет вперед.

2.3 Искусственный интеллект в распознавании и синтезе речи

На 2023 год рынок искусственного интеллекта в распознавании и синтезе речи растёт, как в мире, так и в России. Это обусловлено низкой стоимостью устройств распознавания речи, осведомленность о преимуществах использования этой технологии и технологические достижения способствуют росту рынка распознавания речи. Среди всех рынков стоит выделить Китай с ростом рынка в 21,3 % и США 13,9 % на 2022 год [252].

Объём мирового рынка распознавания голоса и речи в 2022 году оценивался в 17,17 млрд долларов США, и ожидается, что к 2030 году объём рынка составит 48,8 млрд долларов США при среднегодовом росте примерно на 21,30 % за рассматриваемый период [252].

Ожидается, что рынок будет определяться технологическими достижениями и растущим внедрением передовых электронных устройств. Активируемые голосом биометрические данные, используемые в целях безопасности, помогают в предоставлении доступа аутентифицированным пользователям для выполнения транзакции. Растущее использование голосовой биометрии является одним из основных факторов, стимулирующих рост рынка. Растущий спрос на голосовые навигационные системы и рабочие станции стимулирует рост в сегментах аппаратного и программного обеспечения.

Рынок распознавания речи демонстрирует быстрый рост доходов из-за расширения использования в секторах образования по всему миру. Кроме того, растущий спрос на точные и простые в использовании API-интерфейсы распознавания речи, подходящие для различных областей и языков, способствует росту рынка. Индустрия преобразования речи в текст, которая воспринимает быструю оцифровку, положительно влияет на доли рынка.

Вспышка пандемии Covid-19 положительно повлияла на рынки речи и распознавания голоса. Поскольку большая часть населения работает из дома, спрос на дистанционное обслуживание заметно увеличился, что стало причиной высокой нагрузки на call-центры. Данный фактор стал причиной бурного развития рынка искусственного интеллекта в распознавании и синтезе речи.

В настоящее время можно выделить следующие компании, занимающие значительную часть рынка в распознавании и синтезе речи в мире: Google Inc., Iflytek Co. Ltd., Microsoft Corp., Advanced Voice Recognition Systems Inc., Apple Inc., Amazon.com Inc., IBM, Speak2web, Facebook Inc, Rephrase.ai.

Корпорация Google известна следующими продуктами в области распознавании и синтезе речи:

- Google Assistant – виртуальный помощник на базе искусственного интеллекта, который поддерживает распознавание голоса и понимание естественного языка для выполнения задач и ответов на вопросы;

- Google Speech-to-Text – служба, которая использует передовые алгоритмы нейронной сети для преобразования устной речи в письменный текст, что позволяет использовать такие приложения, как службы транскрипции и голосовые команды;

- Google Text-to-Speech – служба, которая синтезирует текст в речь, позволяя приложениям читать текст вслух или создавать аудиоконтент из письменного текста;

- Google Cloud Speech API – облачный API, предоставляющий разработчикам службы распознавания речи для интеграции в свои приложения;

- Google Voice – служба телефонии, которая предлагает расшифровку голосовой почты, позволяя пользователям читать свои сообщения голосовой почты в виде текста;

- Google Meet – платформа для видеоконференций, которая предлагает субтитры и транскрипцию в реальном времени, что делает ее более доступной для пользователей с нарушениями слуха;

- Google Translate – служба языкового перевода, поддерживающая голосовой ввод для перевода произносимых слов и фраз на другие языки;

- Google Home – линейка умных динамиков, которые используют Google Assistant для ответа на голосовые команды, ответов на вопросы и управления устройствами умного

дома;

— Google Gboard – приложение–клавиатура для устройств Android и iOS, включающее функцию голосового набора, позволяющую пользователям диктовать текст, а не печатать.

Iflytek Co. Ltd – китайская технологическая компания, специализирующаяся в области искусственного интеллекта, распознавания и синтеза речи, голосовых помощников. Рассмотрим основные продукты в области распознавании и синтезе речи:

— Платформа iFLYTEK: облачная платформа, предлагающая различные услуги распознавания и синтеза речи, такие как ASR (автоматическое распознавание речи), TTS (преобразование текста в речь) и распознавание голосовых отпечатков;

— iFLYTEK Super Brain: система искусственного интеллекта для обучения, обработки естественного языка и технологий распознавания речи, позволяющая понимать человеческую речь и реагировать на нее в режиме реального времени;

— iFLYTEK Voice Assistant: виртуальный помощник, который использует передовые технологии распознавания речи и обработки естественного языка для предоставления голосовой помощи в различных приложениях, таких как интеллектуальные устройства, домашняя автоматизация и поддержка клиентов;

— iFLYTEK Translator: устройство для перевода в режиме реального времени, поддерживающее несколько языков и использующее технологии распознавания и синтеза речи для обеспечения мгновенного перевода между разными языками;

— Интеллектуальный голосовой робот iFLYTEK: голосовой робот, который может участвовать в разговорах, отвечать на вопросы и предоставлять информацию, используя технологии распознавания и синтеза речи;

— Система оценки речи iFLYTEK: решение для оценки и улучшения качества речи путем анализа разговорной речи, оценки произношения и предоставления отзывов об областях, требующих улучшения;

— Распознавание голосов iFLYTEK: система биометрической аутентификации, использующая технологию распознавания голосовых отпечатков для идентификации и проверки личности человека на основе его уникальных голосовых характеристик;

— iFLYTEK Text-to-Speech Engine: программное обеспечение, которое преобразует письменный текст в естественно звучащую речь, поддерживает несколько языков и используется в различных приложениях, таких как электронные книги, навигационные системы и голосовые помощники;

— Автоматического распознавания речи iFLYTEK: программное обеспечение, которое преобразует разговорный язык в письменный текст, позволяя приложениям с

голосовым управлением и службам транскрипции.

Основными продуктами корпорации Microsoft в области распознавания и синтеза речи являются:

— Cortana – виртуальный помощник Microsoft, использующий распознавание речи для ответа на голосовые команды и вопросы.

— Распознавание речи Windows – встроенное программное обеспечение для Windows, которое позволяет пользователям управлять своим компьютером и диктовать текст с помощью голосовых команд.

— Speech API – комплект для разработки программного обеспечения для добавления в приложения возможностей распознавания и синтеза речи;

— Microsoft Translator – служба перевода, которая использует распознавание и синтез речи для перевода произносимых слов с одного языка на другой;

— Speech Platform – набор инструментов и технологий для разработки приложений с поддержкой речи, включая механизмы распознавания и синтеза речи;

— Speech Services – облачные API-интерфейсы, которые предлагают разработчикам преобразование речи в текст, преобразование текста в речь и другие функции, связанные с речью.

Advanced Voice Recognition Systems – это компания, занимающаяся разработкой продуктов и решений в области распознавания и синтеза речи. Продуктами компании являются:

— AVRSpeechKit: комплект для разработки программного обеспечения для добавления расширенных возможностей распознавания и синтеза речи в мобильные и настольные приложения.

— AVRSpeaker: программное обеспечение для синтеза речи, которое позволяет создавать реалистичные и естественно звучащие компьютерные голоса.

— AVRCommander: программное обеспечение для распознавания голоса, которое позволяет пользователям управлять своими устройствами с помощью голосовых команд.

— AVRSpeechAnalyzer: программный инструмент, который позволяет анализировать и расшифровывать аудиозаписи, что делает его полезным для изучения речи и языка.

— AVRDictate: программное обеспечение для распознавания голоса, которое позволяет пользователям диктовать текст непосредственно на свои компьютеры, избавляя от необходимости ручного набора текста.

— AVRLearn: программное обеспечение для распознавания и синтеза речи, которое обеспечивает интерактивный опыт изучения языка для учащихся всех возрастов и

уровней.

Apple – американская корпорация, разработчик персональных и планшетных компьютеров, аудиоплееров, смартфонов, программного обеспечения и цифрового контента. В контексте распознавания и синтеза речи, Apple разработала:

- Siri – виртуальный помощник Apple, который использует распознавание речи для ответа на голосовые команды и запросы;

- Dictation – встроенная функция преобразования речи в текст в macOS, которая позволяет пользователям диктовать текст в любое приложение;

- VoiceOver – технология чтения с экрана, обеспечивающая синтез речи для слабовидящих пользователей для доступа к своим устройствам и навигации по ним;

- Speak Screen – функция iOS, которая зачитывает пользователям вслух содержимое экрана;

- Speech Synthesis API – программный интерфейс в macOS, который позволяет разработчикам добавлять возможности синтеза речи в свои приложения;

- Core ML – платформа машинного обучения на iOS, которую можно использовать для построения моделей распознавания и синтеза речи.

Amazon.com – американская многонациональная технологическая компания, специализирующаяся на электронной коммерции, облачных вычислениях, онлайн-рекламе и искусственный интеллект. Также Amazon.com разрабатывает технологии голосового управления и голосовых ассистентов, например:

- Amazon Alexa: персональный помощник с голосовым управлением, который использует распознавание и синтез речи, чтобы отвечать на вопросы, воспроизводить музыку, управлять «умными» устройствами;

- Amazon Polly: сервис преобразования текста в речь, использующий передовые технологии глубокого обучения для синтеза естественно звучащей речи на нескольких языках;

- Amazon Transcribe: служба автоматического распознавания речи, которая преобразует речь в текст в режиме реального времени, позволяя разработчикам создавать приложения для расшифровки аудиозаписей;

- Amazon Comprehend: служба обработки естественного языка, использующая машинное обучение для извлечения информации и настроек из текстовых и речевых данных.

- Amazon Lex: сервис для создания диалоговых интерфейсов с использованием голоса и текста на основе автоматического распознавания речи и понимания естественного языка.

- Amazon Transcribe Medical: служба автоматического распознавания речи,

разработанная специально для поставщиков медицинских услуг и позволяющая им быстрее и точнее расшифровывать медицинские разговоры и медицинские записи.

IBM – это американская компания, которая специализируется на разработке программного обеспечения, облачных технологий, искусственного интеллекта и когнитивных вычислений. В области распознавания и синтезе речи предлагает следующие продукты:

- Watson Speech to Text – облачный сервис распознавания речи, который преобразует аудио и голос в письменный текст;

- Watson Text to Speech – облачный сервис, который преобразует письменный текст в естественно звучащий звук на различных языках и с разными голосами;

- IBM Voice Gateway – программная платформа, позволяющая осуществлять голосовое взаимодействие с корпоративными веб-приложениями и мобильными приложениями;

- IBM Watson Assistant – диалоговая платформа искусственного интеллекта, которая может понимать и отвечать на голосовые команды и запросы пользователя;

- IBM Speech Analytics – программное решение, которое анализирует и извлекает информацию из взаимодействия и разговоров с клиентами;

- IBM SpeechViewer – инструмент, который визуализирует речевые данные и помогает выявлять закономерности и тенденции в устном содержании.

Rephrase.ai – индийская компания, работающая в области синтеза речи. Основным продуктом компании является инструмент обработки естественного языка, который генерирует человеческую речь из письменного текста. Инструмент использует алгоритмы глубокого обучения для анализа введенного текста и преобразования его в высококачественный аудиовыход на нескольких языках и с разными голосами. Решение Rephrase.ai имеет различные приложения, включая озвучивание видео, виртуальных помощников, изучение языков и аудиокниги. Компания позиционирует свой продукт как более эффективную и экономичную альтернативу актерам озвучивания и студиям звукозаписи.

Объем российского рынка искусственного интеллекта в области распознавания и синтезе речи в 2022 году достиг 23 млрд рублей, а крупнейшей компанией, занимающей около 70 % рынка является – Яндекс [253]. Большую долю объема российского рынка занимают следующие компании: Яндекс, Группа ЦРТ, VK, Amvera, Tinkoff.

Яндекс известна следующими продуктами в области распознавания и синтезе речи:

- Yandex SpeechKit Cloud – набор API-интерфейсов, которые позволяют разработчикам интегрировать функции распознавания и синтеза речи в свои приложения, продукты и услуги;

— Яндекс.Переводчик – онлайн–сервис языкового перевода, который использует технологию распознавания речи для определения разговорного языка и обеспечивает автоматический перевод;

— Алиса Голосовой помощник – виртуальный помощник Яндекса, который использует распознавание речи и голосовые команды на естественном языке и помогает в различных задачах;

— Яндекс.Навигатор – мобильное навигационное приложение, которое предлагает голосовые указания на основе информации о дорожном движении в реальном времени, включая распознавание речи для простого голосового управления;

— Яндекс.Музыка – служба потоковой передачи музыки, которая позволяет пользователям искать и находить музыку с помощью голосовых команд;

— Яндекс.Браузер – веб–браузер со встроенными функциями голосового поиска, позволяющий пользователям выполнять поиск в Интернете и управлять определенными функциями браузера с помощью голосовых команд;

— Яндекс.Диск – сервис облачного хранилища, поддерживающий распознавание голоса, позволяющий пользователям искать файлы и перемещаться по интерфейсу с помощью голосовых команд;

— Яндекс: Умный дом – Платформа для управления устройствами умного дома с помощью голосовых команд с использованием технологий распознавания и синтеза речи Яндекса через Алису;

— Яндекс.Станция – умная колонка, которая интегрируется с экосистемой Яндекса и предлагает голосовое управление различными сервисами и устройствами;

— Яндекс.Маркет – платформа онлайн–покупок, предлагающая возможности голосового поиска, позволяющие покупателям находить и сравнивать товары, используя запросы на естественном языке.

Группа ЦРТ – российская компания, занимающаяся разработкой инновационных систем в сфере распознавания лиц, голосовой биометрии, распознавания и синтеза речи, многоканальной записи, обработки и анализа аудио– и видеоинформации. Компания известна следующими продуктами в области распознавания и синтезе речи:

— VoiceKey – это голосовое биометрическое решение, которое обеспечивает безопасную аутентификацию и идентификацию пользователей на основе их уникальных голосовых шаблонов;

— GridID – это решение для мультимодальной биометрической аутентификации, сочетающее в себе технологии распознавания голоса и лица;

— VoiceGrid™ RT – это технология распознавания голоса в реальном времени,

которую можно использовать для аутентификации голоса и преобразования речи в текст;

— VoiceGrid™ X – это система идентификации говорящего, которая может анализировать и сравнивать большое количество голосовых данных практически в режиме реального времени.

VK – российская инвестиционная технологическая корпорация, разработавшая следующие технологии в области распознавания и синтезе речи:

— Голосовой помощник Маруся – голосовой помощник VK, предназначенный для помощи в таких функциях, как воспроизведение музыки, управление устройствами умного дома и ответы на вопросы;

— VK Капсула – умная колонка со встроенным голосовым помощником «Маруся»;

— VK Assistant – это голосовой помощник, который умеет обрабатывать естественный язык для того, чтобы пользователь мог задавать ему вопросы и получать ответы;

— Платформа для создания новостных видео – сервис для создания новостных и репортажных видео студийного качества.

Amvera – компания разработчик сервисов и инфраструктуры для программистов и IT-компаний. Основным продуктом компании в области распознавания и синтезе речи является Amvera Speech. Amvera Speech – платформа для распознавание спонтанной русской речи, которую используют крупные промышленные компании как Северсталь и Актон.

Основным продуктом Tinkoff в области распознавания и синтезе речи является Tinkoff VoiceKit – сервис распознавания и синтеза речи для создания голосовых роботов, озвучки видео, перевода аудио в текст и текста в аудио.

Центр компетенций НТИ МФТИ - это платформа, созданная в рамках национальной технологической инициативы (НТИ) и функционирующая на базе Московского физико-технического института (МФТИ). Его задача заключается в развитии инновационных проектов и технологий в области искусственного интеллекта, квантовых технологий, нанотехнологий, биотехнологий, космических технологий и других областей. Центр также обучает специалистов в этих областях и участвует в международных научных и технологически проектах. Основным проектом центра компетенций НТИ в области речевых технологий являются речевая аналитика и голосовые роботы. Создание специализированной системы распознавания и анализа речи на основе методов компьютерной лингвистики и нейросетевых технологий, в том числе с применением процессоров Эльбрус и NeuroMatrix.

Потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно-технической деятельности в Российской Федерации в части искусственного интеллекта распознавании и синтезе речи являются государственные органы и ведомства, исследовательские и научно-образовательные

учреждения, разработчики робототехники и автономных систем, программного обеспечения в области машинного обучения, крупные производственные предприятия, крупные банки и электронная и розничная торговля. Из заинтересованных компаний можно выделить: Газпром Нефть, X5 Group, Северсталь, ПИК, МТС и др.

2.4 Искусственный интеллект в рекомендательных системах и интеллектуальных системах поддержки принятия решений

На 2023 год рынок искусственного интеллекта в рекомендательных системах и интеллектуальных системах поддержки принятия решений растет, как в мире, так и в России. Прогнозируется, что объём мирового рынка систем рекомендаций достигнет 54 млрд долларов США к 2030 году по сравнению с 4,11 млрд долларов США в 2022 году, и ожидается, что рост составит 37 % в течение прогнозируемого периода до 2030 года [254].

Рост рынка искусственного интеллекта в РС и СППР можно объяснить следующими причинами:

— Рост электронной коммерции и персонализации. Индустрия электронной коммерции выросла в геометрической прогрессии, и клиенты ожидают персонализированного опыта при совершении покупок в Интернете. Системы рекомендаций позволяют компаниям анализировать предпочтения клиентов, их поведение и историю покупок, чтобы предоставлять персонализированные предложения и рекомендации;

— Достижения в области искусственного интеллекта и машинного обучения. За последние несколько лет технологии ИИ и МО значительно улучшились, что позволяет СППР и рекомендательным системам принимать более качественные и точные решения;

— Необходимость улучшения процесса принятия решений и повышения эффективности. Поскольку предприятия работают в условиях растущей конкуренции, решающее значение имеет быстрое и информированное принятие решений. Интеллектуальные системы поддержки принятия решений и системы рекомендаций помогают организациям принимать более быстрые и разумные решения, предоставляя информацию на основе данных, что приводит к повышению эффективности и прибыльности;

— Растущий спрос на информацию о клиентах, основанную на данных. Предприятия больше внимания уделяют пониманию предпочтений, привычек и тенденций клиентов, чтобы оставаться впереди на рынке. Рассматриваемая технология позволяет компаниям принимать решения на основе собранной информации о клиентах, что позволяет создавать более успешные маркетинговые стратегии, а также востребованные продукты или услуги;

— Снижение затрат и повышение рентабельности инвестиций. Внедрение СППР и

систем рекомендаций может привести к снижению затрат и повышению рентабельности инвестиций для предприятий за счет автоматизации процессов принятия решений и сокращения ручного вмешательства. Возможность рационализировать бизнес–процессы и улучшить операционную деятельность побуждает предприятия инвестировать в эти системы, способствуя дальнейшему росту рынка.

Также важно заметить влияние пандемии COVID–19, которая стала одним из главных факторов бурного развития электронной торговли и электронных услуг в 2019–2021 годах, как показал опыт многих компаний – рекомендательные системы являются гарантией соответствия потребностей потребителей и продуктов, что несет в себе удовлетворенность клиентов и их удержание. Данные системы уже проявили себя как отличный инструмент для улучшения обслуживания клиентов, которые, вероятно, будут способствовать расширению рынка.

В настоящее время можно выделить следующие компании, занимающих значительную часть рынка искусственного интеллекта в РС и СППР в мире: IBM Corporation, Google LLC (Alphabet Inc), Amazon Web Services Inc, SAS, SAP, Alibaba Group, Baidu, Netflix, Microsoft Corporation, Salesforce.com Inc, Unbxд Inc, Oracle Corporation, Intel Corporation, SAP SE, Hewlett Packard Enterprise Co., Qubit Digital, Algonomy Software Pvt Ltd, Recolize GmbH, Adobe Inc и др.

Компания SAS является крупнейшей в мире частной ИТ–компанией, специализирующейся на разработке и продаже решений и услуг в области бизнес–аналитики. SAS является пионером в разработке передовых систем рекомендаций и интеллектуальных систем поддержки принятия решений на основе искусственного интеллекта и машинного обучения. Последние достижений и продукты компании в рассматриваемое области:

— SAS Intelligent Decisioning. Платформа помогает организациям автоматизировать, оптимизировать и координировать бизнес–решения с помощью расширенной аналитики и искусственного интеллекта. Решение предназначено для создания высоко персонализированного клиентского опыта, который способствует вовлечению и лояльности;

— SAS Customer Intelligence – этот продукт обеспечивает аналитику клиентов, включая сегментацию и персонализированные рекомендации, основанные на поведении клиентов;

— SAS Enterprise Miner – этот продукт предлагает возможности интеллектуального анализа данных и прогнозного моделирования для создания персонализированных рекомендаций;

— SAS Viya – эта платформа предоставляет возможности машинного обучения для

создания интеллектуальных систем поддержки принятия решений;

— SAS Visual Data Mining and Machine Learning – этот продукт предлагает комплексный анализ данных и прогнозное моделирование для создания персонализированных рекомендаций и систем поддержки принятия решений;

— SAS Decision Manager – этот продукт автоматизирует процессы принятия решений с помощью бизнес-правил и моделей прогнозирования;

— SAS Fraud Management – этот продукт использует машинное обучение и аналитику для обнаружения и предотвращения мошеннических действий;

— SAS Intelligent Advertising for Publishers – этот продукт помогает издателям получать доход, предоставляя пользователям персонализированную рекламу;

— SAS Model Manager – этот продукт помогает создавать и развертывать прогностические модели для систем рекомендаций и поддержки принятия решений;

— SAS Visual Analytics – этот продукт предоставляет интерактивные диаграммы, графики и другие визуальные элементы, помогающие принимать обоснованные решения;

— Рекомендательные системы на основе глубокого обучения. Данные системы способны помочь предприятиям персонализировать взаимодействие с клиентами. Системы используют передовые алгоритмы для анализа поведения, предпочтений и взаимодействий клиентов, чтобы рекомендовать наиболее подходящие продукты, услуги и предложения;

— Модели прогнозной аналитики для персонализированного маркетинга и взаимодействия с клиентами. Данные модели используют методы машинного обучения для определения предпочтений и поведения клиентов, которые затем можно использовать для персонализации маркетинговых сообщений и предложений.

SAP – это международная компания–разработчик программного обеспечения для менеджмента бизнеса и управления корпоративными процессами. Рассмотрим наиболее популярные продукты компании:

— SAP Predictive Analytics – это сервис обеспечивает возможности прогнозного моделирования и машинного обучения для принятия решений на основе данных и предоставления персонализированных рекомендаций;

— SAP HANA – эта платформа управления данными и аналитики в оперативной памяти позволяет принимать решения на основе данных в режиме реального времени и выполнять расширенную аналитику;

— SAP Customer Activity Repository – это решение собирает данные из различных источников и применяет аналитику для предоставления рекомендаций по повышению вовлеченности и лояльности клиентов;

— SAP Marketing Cloud – эта платформа автоматизации маркетинга использует

прогнозную аналитику и машинное обучение для предоставления персонализированных рекомендаций, и информации для целевого маркетинга;

— SAP Sales Cloud – это решение сочетает в себе функции управления взаимоотношениями с клиентами с машинным обучением для предоставления интеллектуальных рекомендаций отделам продаж;

— SAP Product Content Hub – это облачное решение обеспечивает централизованное управление и распространение информации о продуктах, используя машинное обучение для создания контента и рекомендаций.

— SAP Fieldglass – это решение для управления персоналом использует прогнозную аналитику для оптимизации предложения и спроса на таланты, предоставляя интеллектуальные рекомендации и поддержку принятия решений;

— SAP Supply Chain Execution – это решение для управления логистикой обеспечивает видимость и аналитику в режиме реального времени для принятия решений на основе данных, включая интеллектуальные рекомендации по оптимизации цепочки поставок;

— SAP SuccessFactors – это программное обеспечение для управления персоналом использует аналитику данных и машинное обучение для предоставления рекомендаций по управлению талантами и планированию рабочей силы.

Alibaba Group – это ведущая технологическая компания Китая, разработавшая широкий спектр продуктов и услуг в области рекомендательных систем и интеллектуальных систем поддержки принятия решений. Компания вложила значительные средства в технологии машинного обучения и искусственного интеллекта для создания передовых механизмов рекомендаций, инструментов визуализации данных, интеллектуальных маркетинговых платформ и логистических систем. Основными решениями компании в рассматриваемой области:

— Интеллектуальная служба рекомендаций «Alibaba Cloud» – это облачная служба, использующая алгоритмы машинного обучения для предоставления персонализированных рекомендаций по продуктам, услугам и контенту;

— Alimama – интегрированная маркетинговая платформа, которая использует большие данные и технологии искусственного интеллекта для оптимизации эффективности рекламы и предоставления узконаправленной рекламы потребителям;

— Alibaba Group AI Labs – это научно-исследовательский центр, который разрабатывает передовые технологии искусственного интеллекта в области обработки естественного языка, компьютерного зрения и машинного обучения;

— Youku – это ведущая онлайн-видеоплатформа Китая, которая использует рекомендации контента на основе искусственного интеллекта для персонализации просмотра

пользователями;

— Руководство по покупкам Taobao: эта функция на платформе электронной коммерции Alibaba предоставляет пользователям персонализированные рекомендации по продуктам на основе их истории поиска и покупок;

— Маркетинговая платформа Alibaba Group – эта платформа помогает компаниям использовать большие данные и технологии искусственного интеллекта для улучшения своих маркетинговых кампаний, и увеличения продаж;

— Cainiao: интеллектуальная логистическая платформа, использующая искусственный интеллект для оптимизации операций цепочки поставок и повышения эффективности доставки.

IBM Corporation – многонациональная технологическая компания, предлагающая широкий спектр продуктов и услуг в области рекомендательных систем и интеллектуальных систем поддержки принятия решений. IBM предлагает ряд продуктов, связанных с гендерными системами и интеллектуальными системами поддержки принятия решений, включая такие платформы, как IBM Watson Studio, IBM Watson Marketing и IBM Watson Commerce Insights, которые используют ИИ для предоставления персонализированных рекомендаций клиентам на основе их поведения и предпочтений. Кроме того, IBM также предоставляет инструменты бизнес-аналитики, такие как IBM SPSS Statistics и IBM Predictive Analytics, которые позволяют организациям принимать решения на основе данных на основе информации, полученной из больших и сложных наборов данных. В целом корпорация IBM играет важную роль в продвижении инноваций в области рекомендательных систем и интеллектуальных систем поддержки принятия решений.

Google – это технологическая компания, которая специализируется на интернет-услугах и продуктах, включая веб-поиск, технологии интернет-рекламы, облачные вычисления, программное и аппаратное обеспечение. Одним из примеров использования Google рекомендательных систем является Google Search. Google представляет результаты поисковой системы на основе истории поиска, местоположения и истории просмотров пользователя. Он также использует алгоритмы машинного обучения, чтобы рекомендовать похожие поиски на основе поискового запроса пользователя. Другой пример – Google Maps, где система предоставляет персонализированные рекомендации по ресторанам, отелям и другим достопримечательностям на основе местоположения пользователя и истории поиска. Алгоритмы совместной фильтрации Google также используются в ряде продуктов, включая YouTube, Google Play и Google Ads, для предоставления персонализированных рекомендаций пользователям.

Кроме того, интеллектуальные системы поддержки принятия решений Google, такие

как Google Analytics и Google Cloud AI, предоставляют компаниям информацию и рекомендации, которые позволяют им принимать решения на основе данных, которые улучшают их операции, маркетинг и продажи. В целом работа Google по разработке интеллектуальных систем для рекомендаций и принятия решений оказала значительное влияние на область искусственного интеллекта и изменила то, как мы используем Интернет.

Amazon – глобальная компания электронной коммерции, которая предоставляет широкий спектр продуктов и услуг через свою онлайн-платформу. В области рекомендательных систем и интеллектуальных систем поддержки принятия решений Amazon известна сложными алгоритмами, которые она использует для рекомендации продуктов своим клиентам. Компания использует такие методы, как совместная фильтрация, фильтрация на основе контента и гибридная фильтрация, чтобы предоставлять персонализированные рекомендации по продуктам отдельным клиентам на основе их поведения при просмотре и покупке. Кроме того, Amazon использует искусственный интеллект и машинное обучение для оптимизации своей цепочки поставок и управления запасами, что позволяет компании поставлять продукты более эффективно и результативно. В целом Amazon – компания-первопроходец в области рекомендательных систем и интеллектуальных систем поддержки принятия решений, и ее инновации помогли сформировать отношение бизнеса к этим технологиям.

Microsoft – технологическая компания, предлагающая различные продукты и услуги, включая операционные системы, инструменты повышения производительности, игровые консоли и решения для облачных вычислений. В области рекомендательных систем Microsoft разработала различные решения, такие как Personalizer API, который использует алгоритмы машинного обучения для предоставления пользователям персонализированных рекомендаций. Он также предлагает Azure Machine Learning, облачную службу, которая позволяет предприятиям разрабатывать и развертывать интеллектуальные системы поддержки принятия решений.

Netflix – известная платформа для потоковой передачи контента, которая использует рекомендательные системы и интеллектуальные системы поддержки принятия решений для повышения удобства пользователей на своей платформе. Бизнес-модель компании основана на вовлечении пользователей с помощью персонализированных рекомендаций по контенту. Алгоритм Netflix отслеживает и анализирует историю просмотров пользователей, поисковые запросы и другие данные, чтобы создавать индивидуальные предложения фильмов и телепередач, которые им может понравиться. Это включает в себя использование машинного обучения и искусственного интеллекта для точного прогнозирования предпочтений пользователя. Со временем алгоритм улучшается, поскольку он получает больше данных и

отзывов от пользователей. Таким образом, Netflix использует интеллектуальные системы поддержки принятия решений, чтобы рекомендовать персонализированный контент своим пользователям, что является одной из ключевых функций, которые делают его сегодня одним из самых популярных потоковых сервисов в мире.

Salesforce – это компания, которая предоставляет облачные программные решения для управления взаимоотношениями с клиентами для предприятий любого размера. Что касается рекомендательных систем и интеллектуальных систем поддержки принятия решений, Salesforce предлагает набор инструментов на основе искусственного интеллекта, которые помогают специалистам по продажам, маркетингу и обслуживанию клиентов принимать решения на основе данных и персонализированные рекомендации.

Инструменты Salesforce на базе искусственного интеллекта, в том числе Einstein Analytics, Einstein Prediction Builder и Einstein Next Best Action, используют алгоритмы машинного обучения для обработки больших объёмов данных о клиентах, чтобы выявить тенденции и закономерности, которые могут помочь в принятии решений и направить тактику продаж. Например, Einstein Next Best Action использует данные клиентов для предоставления персонализированных рекомендаций специалистам по продажам и обслуживанию, помогая им предлагать нужные продукты и услуги клиентам в нужное время. В целом, набор инструментов Salesforce на базе искусственного интеллекта помогает компаниям оптимизировать свои усилия в области продаж и маркетинга, улучшать отношения с клиентами и принимать обоснованные решения на основе анализа данных.

Объём российского рынка искусственного интеллекта в области рекомендательных системах и интеллектуальных системах поддержки принятия решений в 2022 году достиг 52 млрд рублей [255]. Большую долю объёма российского рынка занимают следующие компании: Яндекс, VK, Лаборатория Касперского, Сбербанк, Озон, МТС, X5 Retail Group, Цифра, ЛАНИТ, Лига Цифровой Экономики, Glowbyte, Крок, PARMA Technologies Group, Корус Консалтинг, РДТЕХ, Сапран, Бизнесавтоматика, Информационные технологии будущего, Форсайт, Sapiens solutions, Innodata, Navicon, БАРС Групп.

Яндекс – это российская компания, которая активно развивает технологии и продукты в области рекомендательных систем и интеллектуальных систем поддержки принятия решений. Она использует алгоритмы машинного обучения и искусственного интеллекта для создания персонализированных рекомендаций и предсказаний на базе больших данных, включая поисковые запросы, данные о пользовательском поведении и другие данные. Рассмотрим основные продукты компании в данной области:

— Яндекс.Поиск – поисковая система с персонализированными рекомендациями. Алгоритмы машинного обучения дают возможность подбирать самую релевантную

информацию и обеспечивают оптимальный поиск для каждого пользователя.;

— Яндекс.Карты – сервисы навигации и маршрутизации с рекомендациями о точках интереса;

— Яндекс.Такси – сервис заказа такси с использованием интеллектуальных решений для оптимизации процесса оформления заказа и поиска автомобиля;

— Яндекс.Маркет – это онлайн–маркетплейс, который предоставляет пользователям возможность искать и сравнивать товары и услуги. Рекомендательные системы используются в Яндекс.Маркете для того, чтобы подобрать товары, которые могут представлять интерес для пользователя и наилучшим образом соответствовать его потребностям.

— Яндекс.Метрика – веб–аналитический сервис, который позволяет собирать и анализировать данные о поведении пользователей на сайтах и предлагать рекомендации по улучшению пользовательского опыта;

— Яндекс.Музыка – музыкальный сервис с интеллектуальной системой подбора музыки, основывающейся на предпочтениях и поведении пользователей. Когда пользователь слушает музыку, система анализирует его профиль, жанры и предыдущие прослушивания, чтобы порекомендовать ему новые композиции и исполнителей;

— Яндекс.Диск – облачное хранилище данных с автоматической классификацией файлов и предлагаемыми рекомендациями хранения;

— Яндекс.Курьер – сервис доставки с использованием алгоритмов для оптимизации процесса доставки и рекомендаций партнеров для максимальной эффективности.

Лига Цифровой Экономики – это российская компания, занимающаяся разработкой и внедрением интеллектуальных систем поддержки принятия решений и рекомендательных систем в различных сферах деятельности, таких как госсектор, торговля и питание, финансы, телеком и медиа, промышленность. Она помогает своим клиентам использовать данные с помощью алгоритмов машинного обучения и искусственного интеллекта для повышения эффективности бизнеса и принятия наиболее обоснованных решений. Компания предлагает следующие услуги: система анализа и визуализации данных «ALMAX BI», система для построения полномасштабного процесса современного целевого маркетинга «CAMPAIGN MANAGEMENT CVM-BOX», система управления пропускным режимом и мониторинга «СУПРiМ», интеллектуальная платформа «МАЯК BI» и др.

Компания VK использует рекомендательные системы и интеллектуальные системы поддержки принятия решений для различных проектов и сервисов, которые она предлагает. Например, она использует алгоритмы машинного обучения для персонализации почтовых

ящиков своих пользователей, контекстной рекламы на порталах и социальных сетях, а также для рекомендации контента на своих медиа-ресурсах. Это позволяет улучшить взаимодействие пользователей с сервисами VK, предоставляя им наиболее релевантный контент и повышая лояльность к компании. Кроме того, VK также разрабатывает и продвигает рекомендательные системы для других компаний и бизнесов. Рассмотрим наиболее популярные продукты компании использующие рекомендательные системы:

— ВКонтакте – платформа социальной сети, использующая рекомендательную систему, чтобы предлагать пользователям друзей и контент на основе их интересов и вовлеченности;

— Mail.ru – поставщик услуг электронной почты, использующий интеллектуальную систему поддержки принятия решений для фильтрации спама и определения приоритетности важных писем;

— Youla – онлайн-рынок, использующий рекомендательную систему, чтобы предлагать продукты пользователям на основе их истории просмотров и покупок;

— Delivery Club – платформа доставки еды, которая использует рекомендательную систему, чтобы предлагать рестораны и блюда пользователям в зависимости от их местоположения и предпочтений;

— Maps.me – навигационное приложение, использующее интеллектуальную систему поддержки принятия решений для предоставления пользователям персонализированных указаний в зависимости от их вида транспорта и условий движения.

— Одноклассники – социальная сеть, которая использует рекомендательную систему, чтобы предлагать пользователям группы, события, обсуждения на основе их личных связей и интересов;

— TamTam – приложение-мессенджер, в котором используется интеллектуальная система поддержки принятия решений для предоставления пользователям быстрых ответов и предлагаемых ответов на основе содержания их сообщений;

— HeadHunter – платформа для поиска работы, которая использует рекомендательную систему, чтобы предлагать релевантные объявления о вакансиях пользователям на основе их навыков и опыта;

— Облако Mail.ru – облачный сервис хранения данных, использующий интеллектуальную систему поддержки принятия решений для организации файлов и предложения связанных документов на основе их содержания и ключевых слов.

Сбербанк первый банк России, который активно использует рекомендательные системы и интеллектуальные системы поддержки принятия решений. В рамках этой области Сбербанк разрабатывает инструменты, которые помогают клиентам быстрее и удобнее

находить необходимые услуги и продукты, а также предоставлять им персональные рекомендации на основе анализа их истории взаимодействия с банком.

Помимо этого, Сбербанк активно использует предсказательную аналитику и системы поддержки принятия решений. Они позволяют банку анализировать поведение клиентов, предсказывать их потребности и предлагать наиболее оптимальные продукты и услуги. Также эти инструменты помогают банку более эффективно управлять финансовыми рисками и принимать обоснованные инвестиционные решения. Рассмотрим некоторые продукты Сбербанка:

- Сервис автоматизированной оценки соискателей при найме консультантов. С помощью искусственного интеллекта банк присваивает скоринг–балл кандидату и вычисляет, скоро ли тот решит уволиться;

- Big Five – технология позволяет составить психологический портрет личности и оценить его благонадёжность для банка на основе пяти черт характера: добросовестность, открытость, общительность, законопослушность и эмоциональная неустойчивость;

- Рекомендательная система для платежей и переводов в приложении Сбербанка – предсказывает покупки клиента и рекомендует ему магазины партнеров. Кроме того, рекомендует повторяющиеся платежи.

Озон – это российская интернет–компания, занимающаяся онлайн–торговлей. Для своего сервиса компания использует следующие инструменты в области систем принятия решений:

- Ozon Recommender System для онлайн–ритейла – эта система предоставляет клиентам персонализированные рекомендации по продуктам на основе их истории покупок и поведения;

- Интеллектуальная система поддержки принятия решений для управления запасами – эта система использует анализ данных, чтобы помочь оптимизировать уровень запасов на складах;

- Интеллектуальная система поддержки принятия решений для управления цепочками поставок. Система помогает оптимизировать операции в цепочках поставок, анализируя данные и предоставляя рекомендации;

- Система обнаружения мошенничества Ozon: эта система использует алгоритмы машинного обучения для выявления потенциального мошенничества и предотвращения его возникновения.

- Интеллектуальная система поддержки принятия решений для маркетинга: эта система помогает компаниям оптимизировать свои маркетинговые кампании, анализируя данные и предоставляя рекомендации.

— Чат-бот Ozon: эта система позволяет предприятиям оказывать поддержку клиентам и отвечать на вопросы клиентов с помощью чат-бота на базе искусственного интеллекта.

ФГАОУ ВО "Национальный исследовательский университет – Высшая школа экономики" занимается межотраслевыми технологиями искусственного интеллекта и искусственный интеллект для иных приоритетных отраслей экономики и социальной сферы. Основными проектами Высшей школы экономики являются:

- Рекомендательный сервис для автоматизации аналитических процессов научно-технической деятельности,
- ИИ в макро моделировании и прогнозировании экономических процессов и финансовых взаимосвязей с учётом сентимента участников рынка,
- Исследование двустороннего рынка в краудсорсинговых системах,
- ИИ в информационных процессах: алгоритмизация выбора, ограничение и ответственное медиапотребление.

Центр компетенций НТИ МФТИ - это платформа, созданная в рамках национальной технологической инициативы (НТИ) и функционирующая на базе Московского физико-технического института (МФТИ). Его задача заключается в развитии инновационных проектов и технологий в области искусственного интеллекта, квантовых технологий, нанотехнологий, биотехнологий, космических технологий и других областей. Центр также обучает специалистов в этих областях и участвует в международных научных и технологически проектах. Основными проектами центра компетенций НТИ являются:

- Технологии оперативного мониторинга природных территорий: разработка систем оперативного мониторинга природных территорий на основе локальной глубокой обработки данных дистанционного зондирования и технического зрения при помощи технологий искусственного интеллекта,
- Экспертные системы для планирования морской сейсморазведки: создание экспертной системы и программно-аппаратного комплекса для планирования морской сейсморазведки в реальном времени,
- Системы поддержки принятия врачебных решений: построение экспертных систем поддержки принятия врачебных решений в области флюорографии, маммографии, кардиографии с использованием поисковых систем и технологий глубокого машинного обучения,
- Система оптимизации движения ж/д транспорта «Прогноз»: создание многофункциональной интеллектуальной системы прогнозирования, моделирования и оптимизации движения железнодорожного транспорта, в том числе с применением

технологий анализа больших данных,

— Техническое зрение в радиодиапазоне: создание системы автоматизированного проектирования радиолокационной техники с применением технологий искусственного интеллекта и машинного обучения для задач распознавания аэродинамических целей по дальностным портретам,

— Мультиагентная платформа интеллектуального трейдинга для самоорганизующихся энергосистем: разработка мультиагентной платформы сервисов интеллектуального трейдинга самоорганизующихся микроэнергетических систем,

— Технологии для интеллектуальной системы управления городской инфраструктурой: разработка аппаратно-программного комплекса автоматизации управления энергоресурсами, безопасностью, с применением беспроводной телеметрической сети типа LPWAN, а также автотранспортом и другими подвижными объектами с применением сетей и технологий V2X, 5G, интеллектуальных приборов и технологий искусственного интеллекта,

— Системы управления процессами разработки месторождений: создание экспертной системы и программного комплекса мониторинга и оптимального управления разработкой месторождений с трудноизвлекаемыми запасами на базе алгоритмов машинного и глубинного обучения,

— Проектирование фармпрепаратов и прогнозирование их свойств: разработка технологий искусственного интеллекта для создания новых лекарственных препаратов и прогнозирования их свойств.

Потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части искусственного интеллекта в рекомендательных системах и интеллектуальных системах поддержки принятия решений являются государственные органы и ведомства, исследовательские и научно–образовательные учреждения, разработчики робототехники и автономных систем, программного обеспечения в области машинного обучения, крупные производственные предприятия, крупные банки и электронная и розничная торговля. Из заинтересованных компаний можно выделить: X5 Group, Роснефть, ПИК, АвтоВаз, КамАЗ.

2.5 Генеративный искусственный интеллект

Объём мирового рынка генеративного ИИ оценивался в 10,2 млрд долларов в 2022 году [256]. Рост генеративного искусственного интеллекта можно объяснить несколькими факторами, включая достижения в области технологий, увеличение инвестиций, доступность больших наборов данных и растущий спрос на приложения ИИ в различных отраслях.

Рассмотрим основные из них:

— Разработка и совершенствование новых алгоритмов, таких как генеративно–состязательные сети и трансформеры, значительно расширили возможности генеративных систем ИИ. Эти новые модели позволили ИИ создавать более реалистичный и творческий контент, включая изображения, видео, иллюстрации, музыку и текст на естественном языке;

— Постоянный рост мощности графических и тензорных процессоров, привод к тому, что исследователи могут обучать более сложные и глубокие модели ИИ;

— Рост Интернета и цифровизация практически всех сфер и услуг упростили доступ к большим объёмам данных и их хранение. Эти наборы данных постоянно расширяются, что позволяет системам ИИ постоянно учиться и повышать свою производительность;

— Успех и потенциальные возможности применения генеративного ИИ привели к увеличению инвестиций как со стороны частных, так и государственных организаций, что в значительной степени способствовало его росту. Различные компании, в том числе технологические гиганты, такие как Google, IBM и Microsoft, вкладывают значительные средства в исследования и разработки в области ИИ;

— Спрос на приложения ИИ. Такие отрасли, как развлечения, реклама, маркетинг, здравоохранение, финансы и другие, все больше полагаются на генеративный ИИ для улучшения своих продуктов и услуг;

— Исследования в области этики и безопасности. По мере того, как ИИ продолжает развиваться, исследователи и организации решают проблемы этики и безопасности, связанные с контентом, созданным ИИ. Эти усилия поощряют ответственное использование и развитие генеративного ИИ, обеспечивая долгосрочный рост и широкое распространение.

Таким образом, по прогнозам аналитиков, объём рынка достигнет 126,5 миллиарда долларов к 2031 году, увеличившись в среднем на 32 % с 2022 по 2031 год [256].

Вспышка COVID–19 оказала положительное влияние на рынок из–за роста потребности предприятий в генеративном ИИ для улучшения потребностей своих клиентов и создания персонализированного опыта, такого как персонализированные музыкальные плейлисты, персонализированные новостные ленты и персонализированные рекомендации по продуктам, стимулирующие рост рынка.

В настоящее время можно выделить следующие компании, занимающие значительную часть рынка генеративного искусственного интеллекта в мире: IBM Corporation, Genie AI Ltd., MOSTLY AI Inc., Google LLC, Rephrase.ai, Amazon Web Services, Inc., Microsoft Corporation, Adobe., Synthesia, Midjourney, OpenAI.

IBM – американская компания, являющаяся лидером в области искусственного интеллекта. Компания проводит множество исследований в области генеративного ИИ, такими как «Deep Generative Models», объединяющие глубокие нейронные сети, генеративные модели, генеративно–сопоставительные сети, моделирование естественного языка и др. Также IBM ряд продуктов связанных с генеративным ИИ:

— Watson Studio – это платформа для обеспечения разработчиков инфраструктурой и инструментами, необходимые для создания и внедрения различных форм генеративных AI–систем.

— IBM Watson Assistant – это платформа чат–ботов с искусственным интеллектом, которая может интерпретировать запросы клиентов и отвечать на них на естественном языке, а также учиться на прошлых взаимодействиях, чтобы улучшить ответы в будущем;

— IBM Watson Discovery – это поисковая система на базе искусственного интеллекта, которая может принимать, анализировать и извлекать данные из неструктурированных источников данных, таких как отраслевые публикации, технические руководства и внутренние документы;

— Watsonx.ai – это платформа для обучения и разработки приложений на искусственном интеллекте. Она позволяет создавать различные AI–сценарии и решения для бизнеса, включая чат–боты, анализ данных, машинное обучение и другие.

Midjourney – исследовательская компания и разрабатываемое ею одноименное программное обеспечение искусственного интеллекта, генерирующее изображения по текстовым описаниям. Midjourney – это программа искусственного интеллекта, которая генерирует изображения на основе описаний на естественном языке, называемых подсказками.

OpenAI – это американская компания, занимающаяся исследованиями в области искусственного интеллекта. OpenAI разработала передовые алгоритмы и модели, которые могут генерировать сложные и творческие результаты, такие как текст, изображения и музыка. Основными продуктами компании являются: GPT–4 и DALL–E.

GPT–3 – одна из самых продвинутых языковых моделей. Это языковая модель ИИ, которая может генерировать человекоподобный текст большого объема и достаточно хорошего качества. GPT–3 использует методы глубокого обучения, в частности архитектуру нейронной сети, называемую преобразователями, для анализа и понимания огромных объемов текстовых данных, а затем генерирует текстовые выходные данные на основе этих данных, такие как статьи, рассказы, стихи и технические записи. Модель имеет емкость до 175 миллиардов параметров и способна понимать и генерировать естественный язык на уровне сложности, почти неотличимом от человеческого письма.

DALL-E является усовершенствованной генеративной моделью искусственного интеллекта, которая создает изображения из текстовых описаний. Он использует архитектуру нейронной сети, называемую генеративно-состязательная сеть, для создания высококачественных изображений, соответствующих текстовым описаниям. DALL-E был обучен на большом наборе данных изображений и может генерировать широкий спектр уникальных и творческих изображений на основе текстового ввода.

Genie AI Ltd – компания, занимающаяся разработкой программного обеспечения, специализирующаяся на генеративном ИИ. Основным продуктом компании является сервис с ассистентом по правовым вопросам. Данный ассистент способен отвечать на юридические вопросы, генерируя ответ, основанный на крупнейшей в мире открытой юридической библиотеке. Компания является первооткрывателем в области мультимодального юридического искусственного интеллекта. Стоит заметить, что искусственный интеллект Genie основан на GPT-4 и собственных проприетарных языковых моделях Genie, разработанных на основе более чем 100 000 юридических документов.

MOSTLY AI Inc – это компания, специализирующаяся на генеративном ИИ, которая занимается созданием искусственных данных, которые затем можно использовать для различных целей, таких как тестирование и улучшение моделей машинного обучения. Продукт компании представляет собой платформу данных на базе искусственного интеллекта, которая позволяет пользователям создавать синтетические данные, которые можно использовать для аналитики, тестирования и обучения моделей машинного обучения. Платформа использует передовые алгоритмы для создания синтетических данных, которые точно отражают поведение и закономерности реальных данных, а также защищают конфиденциальность и анонимность отдельных лиц. Сгенерированные данные можно настроить в соответствии с конкретными наборами данных, что делает их эффективной и экономичной альтернативой сбору и использованию реальных данных.

Google – многонациональная технологическая компания, специализирующаяся на интернет-услугах и продуктах. Google вкладывает значительные средства в исследования и разработку передовых технологий ИИ, включая генеративно-состязательные сети, вариационные автоэнкодеры и другие методы машинного обучения, которые могут создавать новый и реалистичный контент, такой как изображения, видео, музыку и текст.

Одним из наиболее заметных продуктов Google в области генеративного ИИ является алгоритм Deep Dream, который использует нейронную сеть для анализа и обработки изображений, создавая от сюрреалистических до реалистичных произведений искусства. Другим примером является проект Google Magenta, который использует машинное обучение для создания музыки и других форм искусства, поддерживающих человеческое творчество.

Кроме того, Google также интегрировал генеративный ИИ в свои коммерческие продукты, такие как Google Assistant и Google Translate, что позволяет вести более естественный и беглый разговор с пользователями. В целом, инвестиции Google в генеративный ИИ являются свидетельством приверженности компании к расширению границ технологических инноваций и предоставлению своим пользователям новых и захватывающих возможностей.

Rephrase.ai – индийская компания, работающая в области генеративного ИИ, которая фокусируется на разработке инструментов для создания высококачественных видеопрезентаций и анимации. Продукт компании позволяет создавать совершенно новый контент с помощью генеративных алгоритмов, используя фрагменты видео и фотографии. В качестве входного материала можно использовать записи любых голосов, что позволяет генерировать персонализированные видеоролики для веб-сайтов, рекламы, социальных сетей, игр и многое другое. Основными преимуществами и инновациями продукта являются высокое качество создаваемого контента, доступность для широкой аудитории и масштабность. Продукт компании может быть использован как профессиональными в области медиаконтента, так и любителями, которые хотят создавать свои уникальные видеоролики.

Microsoft внедрил в свой браузер Bing технологию генеративной ИИ для улучшения обработки изображений и повышения качества поисковых результатов. В браузере генеративный ИИ помогает повысить качество изображений, связанных с результатами поиска, снижает количество ошибок и повышает точность классификации изображений, что в результате дает более точные и удобные поисковые результаты для пользователей.

Amazon Web Services, Inc. – дочерняя компания Amazon, которая предоставляет платформы облачных вычислений и API-интерфейсы по требованию для частных лиц, компаний и правительств. В области генеративного ИИ AWS предлагает несколько продуктов и сервисов, в том числе Amazon SageMaker – полностью управляемый сервис машинного обучения, который позволяет разработчикам создавать, обучать и развертывать модели машинного обучения в любом масштабе. В частности, SageMaker предлагает инструменты и алгоритмы для разработки и развертывания генеративных моделей, таких как генеративно-состязательные сети и вариационные автоэнкодеры, для таких целей, как генерация изображений и текста. Компания предлагает другие инструменты и сервисы, которые можно использовать в генеративном ИИ, например DeepLens для глубокого обучения на периферийных устройствах и Amazon Rekognition для анализа изображений и видео.

Adobe Inc. – многонациональная компания-разработчик программного обеспечения со штаб-квартирой в США, специализирующаяся на предоставлении программных продуктов для творчества, маркетинга и дизайна потребителям и предприятиям по всему миру. В области генеративного ИИ Adobe инвестирует в разработку инновационных инструментов и

технологий, использующих машинное обучение и искусственный интеллект для создания контента, улучшения дизайна и оптимизации рабочих процессов пользователей.

Одним из таких продуктов, разработанных Adobe, является «Adobe Sensei». Adobe Sensei – это передовая технологическая платформа, обеспечивающая интеллектуальные функции продуктов Adobe. Он использует методы генеративного ИИ для автоматизации рутинных задач, обеспечения контекстуальной осведомленности в рабочих процессах пользователей и улучшения взаимодействия пользователей с программным обеспечением. С помощью этой технологии творческие профессионалы могут быстро создавать художественные активы, предлагать улучшения дизайна и оптимизировать цифровой опыт, используя алгоритмы машинного обучения и нейронные сети глубокого обучения. Adobe Sensei, платформа искусственного интеллекта и машинного обучения, интегрирована практически во всё программное обеспечение: Photoshop, Premiere Pro, Illustrator и др.

Российский рынок генеративного ИИ достаточно молод, что затрудняет точную оценку, также важно заметить, что на момент написания аналитического доклада в открытом доступе отсутствуют аналитические документы по данной отрасли. Можно приблизительно оценить российский рынок относительно мирового, как правило, мировой рынок ИИ в среднем в 10 раз больше рынка отечественного. Таким образом, получим что объём российского рынка, вероятно, может достигать 1,2 млрд. долларов на 2022 год.

Заметными компаниями на рынке, занимающиеся разработкой и инвестициями в генеративный ИИ: Яндекс, Сбербанк, АBBYY, Merlinai, Dasha.AI, DeepPavlov, НИТУ «МИСиС».

Яндекс – российская компания, специализирующаяся на интернет-услугах и продуктах. Главным продуктом компании в области генеративного ИИ является ЯндексGPT, который компания выпустила в 2023 году. ЯндексGPT – это нейронная сеть, которая используется для генерации текстов. В основе ЯндексGPT лежит технология трансформер, разработанная OpenAI. ЯндексGPT обучен на большом количестве текстовых данных и используется для генерации текста в ответ на заданный вопрос или продолжение заданного текста.

Также важным продуктом является Шедевриум – приложение, запущенное в 2023 году, в котором пользователи могут генерировать картинки по текстовому описанию с помощью нейросетей. В конце августа 2022 года «Яндекс» запустил двуязычную версию генератора текстов «Балабоба». Пользователю можно написать одно–два слова на русском или английском языках и выбрать один из стилей – и «Балабоба» создаст осмысленный текст на любую тему, похожий на тексты из интернета, на которых училась модель.

Сбербанк – крупнейший российский банк и один из первооткрывателей в области

генеративного ИИ. Компания разработала нейросеть «Kandinsky 2.1» в 2022 году для генерации изображений по тексту. Модель сильно улучшили и обучили на 170 миллионах связок «текст–изображение». Kandinsky 2.1 содержит 3,3 миллиарда параметров, предыдущая Kandinsky 2.0 – только 2 миллиарда. На 2023 год Сбербанк активно ведет разработку своего чата на основе генеративного ИИ «GigaChat». GigaChat – мультимодальная версия нейросети, которая способна отвечать на вопросы пользователей, поддерживать диалог, создавать тексты и картинки на основе описаний.

Merlinai – российская компания, занимающаяся искусственным интеллектом и машинным обучением. Она разрабатывает и внедряет решения на основе этих технологий для оптимизации бизнес–процессов, повышения эффективности и качества работы систем. Компания специализируется на анализе больших объёмов данных, предиктивном анализе, мониторинге и автоматизации процессов. Она имеет широкий круг клиентов в различных отраслях, таких как финансы, производство, ритейл, здравоохранение и другие. Основным продуктом компании в области генеративного ИИ является нейронная сеть Merlin основанная на GPT–3 и GPT–4, которая способна дать ответ на вопрос, помогать с поиском в интернете, помогать в написании текста, переводить тексты, перефразировать и искать ключевые слова и предложения.

Dasha.AI – российская платформа для быстрой и удобной разработки умных голосовых моделей и чат ботов. Основным продуктом компании являются разговорные модели.

Проект DeepPavlov, разработанный в Лаборатории нейронных систем и глубокого обучения МФТИ, – это библиотека для создания виртуальных ассистентов и анализа текста, построенная на TensorFlow и Keras [257]. Она содержит набор компонентов для быстрого прототипирования диалоговых систем, которые позволяют автоматизировать процессы коммуникации в различных областях деятельности. Платформа обеспечивает полный цикл разработки диалоговых агентов, предназначенных для автоматизации коммуникационных процессов.

НИТУ «МИСиС» в 2022 году разработал генератор пьес «НейроСтаниславский» на основе RuGPT–3. Нейросеть обучалась на произведениях русских классиков – Островского, Чехова, Горького и других. Студенты Университета МИСИС применили прогрессивный подход, при котором искусственный интеллект не заменяет автора, но помогает создавать и совершенствовать сюжет, развивая мысли режиссера.

Потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части генеративного искусственного интеллекта являются компании: промышленные, медицинские, транспортные, энергетические, а также компании

занимающиеся рекламой и маркетингом, дизайном и архитектурой, производством ПО, также важно заметить заинтересованные государственные организации: министерство финансов, здравоохранения, образования и науки и др. Заинтересованным компания являются: Сбербанк, VK, Озон, Медиаскоп, Тинькофф, X5 GROUP, Мегафон и др.

2.6 Автоматическое машинное обучение

По оценкам аналитиков, мировой рынок автоматического машинного обучения достиг 1 млрд. долларов в 2022 году и ожидается, что к 2028 году рынок достигнет размера 6,4 млрд долларов со средним ростом в 44,6 %. Основными причинами роста рынка автоматического машинного обучения являются:

— Увеличение спроса на принятие решений, основанных на данных. Отрасли и предприятия все больше ориентированы на данные, при этом возрастает осознание ценности принятия решений на основе глубокого анализа данных. Решения AutoML позволяют организациям использовать огромное количество данных и быстро получать значимые результаты, даже без глубокого понимания науки о данных;

— Нехватка квалифицированных специалистов по данным. Существует значительный дефицит специалистов по данным на сегодняшнем рынке труда, при этом спрос превышает предложение. AutoML помогает преодолеть этот разрыв, позволяя компаниям автоматизировать задачи машинного обучения и получать результаты без необходимости наличия специалиста–дата-саентиста;

— Сокращение времени и затрат. Платформы AutoML обычно требуют меньше времени и ресурсов для создания моделей по сравнению с традиционными методами машинного обучения, что делает их более экономически выгодными;

— Конкурентное преимущество. Организации, которые используют AutoML, получают конкурентное преимущество перед теми, которые полагаются только на традиционные методы машинного обучения. Это достигается за счет быстрого внедрения инноваций, оптимизации процесса принятия решений и быстрой адаптации к тенденциям рынка;

— Рост облачных сервисов машинного обучения. С увеличением облачных услуг организации могут получить доступ к различным инструментам AutoML без необходимости развертывания выделенной инфраструктуры на месте. Это приводит к снижению затрат и увеличению масштабируемости, что дополнительно стимулирует рост рынка;

— Технологические достижения. Постоянная эволюция алгоритмов машинного обучения, методов предварительной обработки данных и других базовых технологий значительно способствует росту рынка AutoML, так как позволяет создавать сложные и

точные модели для прогнозирования и анализа в различных сферах применения.

Также стоит выделить быстрорастущий спрос AutoML для выявления мошенничества и для персонализированных рекомендаций по продуктам. Что касается отрасли, то в ближайшие годы наибольший рост ожидается в категории здравоохранения. Это связано с ростом спроса на ML в секторе здравоохранения для раннего распознавания заболеваний, обучения, исследований и лечения пациентов быстро и эффективно, при одновременном сокращении денег, времени и ресурсов.

Ключевыми компаниями в области AutoML являются: IBM (США), Oracle (США), Microsoft (США), Google (США), Baidu (Китай), Amazon Web Service (далее AWS) (США), Alteryx (США), Salesforce (США), Altair (США), Teradata (США), H2O.ai (США), DataRobot (США), BigML (США), Databricks (США), Dataiku (Франция), Alibaba Cloud (Китай), Appier (Тайвань), Squark (США), Aible (США), Datafold (США), Boost.ai (Норвегия), Tazi.ai (США), Akkio (США), Valohai (Финляндия), dotData (США), Qlik (США), Mathworks (США), HPE (США) и SparkCognition (США).

IBM – многонациональная технологическая компания, которая разрабатывает и продает различные аппаратные средства, программное обеспечение и промежуточное программное обеспечение. Основными продуктами IBM в области AutoML являются:

- IBM Watson Studio AutoAI – платформа, которая автоматизирует процесс машинного обучения, включая подготовку, разработку функций и создание моделей;
- IBM Watson Machine Learning – платформа, которая позволяет пользователям создавать и развертывать модели машинного обучения;
- IBM Watson OpenScale – платформа, которая устраняет проблемы, связанные с точностью и объективностью модели, за счет функций мониторинга и объяснимости;
- IBM Cloud Pak for Data – платформа, предоставляющая набор инструментов для обработки данных, включая возможности автоматического машинного обучения.

Основными продуктами, разработанными Google в области автоматического машинного обучения, являются:

- AutoML – набор продуктов, использующих алгоритмы поиска нейронной архитектуры для автоматизации процесса построения моделей машинного обучения;
- Cloud AutoML – облачная платформа, которая позволяет компаниям создавать собственные модели машинного обучения без необходимости иметь опыт в области машинного обучения;
- TensorFlow – библиотека машинного обучения с открытым исходным кодом, разработанная Google, которая обеспечивает основу для создания и обучения моделей машинного обучения;

— TensorFlow Extended (TFX) – комплексная платформа для создания и развертывания готовых к работе конвейеров машинного обучения;

— Kubeflow – инструмент машинного обучения с открытым исходным кодом, который упрощает развертывание рабочих процессов машинного обучения и управление ими на платформе Kubernetes.

Baidu – китайская компания, предоставляющая веб-сервисы, основным из которых является поисковая система с таким же названием – лидер среди китайских поисковых систем. Основные продукты Baidu в области автоматического машинного обучения:

— EZDL – веб-платформа, которая позволяет пользователям создавать и обучать модели машинного обучения с помощью интерактивного интерфейса без каких-либо знаний в области программирования;

— PaddlePaddle – платформа глубокого обучения с открытым исходным кодом, которая предоставляет полный набор инструментов для создания и обучения моделей машинного обучения;

— DeepBench – инструмент сравнительного анализа для глубокого обучения, который измеряет производительность аппаратных платформ при выполнении различных задач глубокого обучения;

— FSGAN – модель машинного обучения, которая может создавать реалистичные лица из неполных или некачественных фотографий;

— AI Studio – платформа, предоставляющая набор инструментов для разработки и развертывания моделей машинного обучения в облаке.

Oracle – американская транснациональная корпорация в области компьютерных технологий. Oracle предлагает несколько продуктов в области автоматического машинного обучения:

— Oracle Autonomous Database – это облачная база данных, которая использует алгоритмы машинного обучения для оптимизации производительности, автоматизации задач управления и предотвращения рисков;

— Oracle Data Science – это интегрированный набор инструментов и сервисов, которые позволяют специалистам по обработке и анализу данных, инженерам и аналитикам создавать, развертывать и управлять моделями машинного обучения, совместно работать над проектами и использовать данные в безопасной среде;

— Oracle Machine Learning – это семейство инструментов и технологий, которые позволяют пользователям разрабатывать, развертывать и управлять моделями машинного обучения в экосистеме Oracle. Он предлагает несколько компонентов, в том числе Oracle Machine Learning для Python, Oracle Machine Learning для Spark и Oracle Machine Learning для

SQL.

Основными продуктами компании Amazon Web Service в области автоматического машинного обучения являются:

— Amazon SageMaker – это полностью управляемый сервис, который помогает специалистам по обработке данных и разработчикам быстро создавать, обучать и развертывать модели машинного обучения. SageMaker предоставляет ряд встроенных алгоритмов для распространенных вариантов использования, а также возможность создавать свои собственные алгоритмы;

— AutoGluon – это набор инструментов AutoML с открытым исходным кодом, который можно использовать с AWS. Он автоматизирует процесс обучения и настройки моделей машинного обучения и может использоваться для различных задач, таких как классификация, регрессия и обнаружение объектов;

— Amazon Lookout for Equipment – это сервис, который использует машинное обучение для обнаружения аномалий в промышленном оборудовании, таком как двигатели и насосы. Это может помочь прогнозировать сбои оборудования и оптимизировать графики технического обслуживания;

— Amazon Monitron – это сервис, который использует датчики и машинное обучение для обнаружения ненормальных состояний оборудования и прогнозирования сбоев. Его можно использовать для различного промышленного оборудования, такого как двигатели, насосы и конвейеры;

— AWS DeepRacer – это сервис, который позволяет разработчикам тестировать модели обучения с подкреплением, используя физический гоночный автомобиль. Его можно использовать вместе с Amazon SageMaker для создания и тестирования моделей.

Основными продуктами Microsoft в области автоматического машинного обучения:

— Azure Machine Learning – эта облачная служба предоставляет инструменты и службы для создания, обучения и развертывания моделей машинного обучения. Он включает в себя возможности автоматизированного машинного обучения, которые позволяют пользователям автоматизировать выбор модели и настройку гиперпараметров;

— MLflow – это платформа с открытым исходным кодом помогает управлять жизненным циклом машинного обучения, включая эксперименты, воспроизводимость и развертывание. Он включает в себя компонент автоматизированного машинного обучения, который позволяет пользователям легко обучать и развертывать модели;

— Azure Cognitive Services – набор предварительно созданных API-интерфейсов, которые позволяют разработчикам добавлять в свои приложения интеллектуальные функции, такие как распознавание изображений и речи, понимание языка и анализ настроений;

— Azure Databricks – совместная аналитическая платформа на основе Apache Spark, которую можно использовать для обработки больших данных, машинного обучения и обработки данных;

— Azure AI – набор инструментов и служб, включающий когнитивные службы, средства машинного обучения и предварительно обученные модели, предназначенные для упрощения создания интеллектуальных приложений предприятиями.

Основными продуктами американской компании Alteryx в области автоматического машинного обучения являются Alteryx Designer и Alteryx Analytics Hub. Alteryx Designer – это мощный инструмент для подготовки, смешивания и анализа данных, который оснащен несколькими возможностями построения моделей машинного обучения. Он позволяет пользователям взаимодействовать с данными, создавать модели и с легкостью развертывать их. Alteryx Analytics Hub, с другой стороны, представляет собой облачную платформу, которая предоставляет комплексное решение для специалистов по данным и бизнес-аналитиков для создания, обучения и развертывание моделей машинного обучения. Он спроектирован так, чтобы быть масштабируемым, гибким и простым в использовании, позволяя пользователям создавать настраиваемые рабочие процессы, которые автоматически обновляются по мере поступления новых данных.

Salesforce – американская компания–разработчик программного обеспечения на базе облачных технологий. Основными продуктами Salesforce связанные с областью автоматического машинного обучения:

— Salesforce Einstein – это первый полностью интегрированный ИИ для системы взаимодействия с клиентами. Einstein предоставляет предприятиям возможность применять основные концепции машинного обучения и передовые концепции глубокого обучения ко всем своим процессам и данным, включая данные о клиентах, отрасли и внутренние данные, для непревзойденной точности и передовой автоматизации;

— Einstein Prediction Builder – этот продукт позволяет пользователям создавать собственные прогностические модели без какого-либо опыта программирования;

— Einstein Language – этот продукт позволяет пользователям создавать собственные модели обработки естественного языка для анализа текстовых данных;

— Einstein Vision – этот продукт позволяет пользователям создавать собственные модели классификации изображений для анализа данных изображений;

— Einstein Sentiment – этот продукт позволяет пользователям анализировать настроения и эмоции клиентов в текстовых данных.

Teradata предлагает ряд продуктов для данных и аналитики, включая программное обеспечение, облачные технологии, аппаратное обеспечение, средства управления

экосистемой, приложения и решения с открытым исходным кодом. Основным продуктом в области AutoML является Vantage. Vantage автоматизирует и упрощает весь рабочий процесс машинного обучения, от подготовки данных до развертывания модели, с целью сделать машинное обучение более доступным для неспециалистов. Vantage предлагает ряд функций AutoML, включая оптимизацию гиперпараметров, выбор функций и выбор модели.

DataRobot – платформа машинного обучения для автоматизации, обеспечения и ускорения обработки прогнозной аналитики, которая помогает специалистам по обработке и анализу данных создавать и разворачивать точные прогнозные модели. Среди продуктов компании можно выделить DataRobot AutoML – платформа, которая автоматически создает десятки моделей машинного обучения одним нажатием кнопки, позволяя любому создавать практические модели машинного обучения, независимо от знаний в области науки о данных.

Dataiku – компания по искусственному интеллекту и машинному обучению, основанная в 2013 году. Из всех продуктов стоит выделить Dataiku AutoML – это платформа, которая ускоряет процесс разработки модели с управляемой структурой для многих типов машинного обучения, таких как прогнозирование, кластеризация, прогнозирование временных рядов, задачи компьютерного зрения и многое другое. Также важно осветить продукт MLOps – это инструмент для разработки, развертывания, отслеживания и обслуживания модели машинного обучения на одной платформе.

Прогнозируется, что российский рынок будет демонстрировать быстрый рост за прогнозируемый период. Это связано с растущими расходами на ИТ-инфраструктуру, растущим числом усилий правительства по развитию технологий искусственного интеллекта и плавным экономическим развитием. Российский рынок генеративного автоматизированного машинного обучения достаточно молод, что затрудняет точную оценку, также важно заметить, что на момент написания аналитического доклада в открытом доступе отсутствуют аналитические документы по данной отрасли. Можно приблизительно оценить российский рынок относительно мирового, как правило, мировой рынок ИИ в среднем в 10 раз больше рынка отечественного. Таким образом, объём российского рынка может достигать 0,1 млрд долларов на 2022 год.

В отличие от рассмотренных нами ранее технологий, уже автоматическое машинное обучение являются частью бизнеса. Причем настолько важной и существенной частью бизнеса, что большинство компаний не только не продает свои разработки на сторону, но и предпочитает не распространяться о своих результатах. Заметными компаниями на рынке, занимающиеся разработкой и инвестициями в автоматическое машинное обучение: Яндекс, Сбербанк, Statanly technologies

Основным продуктом VK в области AutoML является «Cloud ML Platform». Cloud ML

Platform – облачная платформа с инструментами и ресурсами для ML/AI-разработки. Целевая версия платформы позволит построить полный цикл работы с машинным обучением: от эксперимента и разработки до запуска готовой версии ML-модели в продакшен.

Лаборатория ИИ «Сбера» (Sber AI Lab) создали фреймворк LightAutoML, сокращённо – LAMA. LAMA – это open-source фреймворк, который позволяет автоматически строить модели обучения с учителем. Изначально решение создавалось для решения следующей задачи: «Чтобы за 15 минут на ноутбуке мы могли создать модель машинного обучения на уровне среднего специалиста, которую потом можно легко передать в продакшен» [258]. Также стоит выделить проект «Сбера» Сбербанк Auto ML, который выпустили в 2019 году [259].

Statanly technologies – российская компания, занимающаяся разработкой интеллектуальных систем, рекомендательных и прогнозных моделей, а также кластеризацией и визуализацией больших данных. Также компания активно создает сервисы AutoML и занимается анализом данных.

Также большой вклад в развитие машинного обучения, нейронных сетей, и ИИ, в том числе сильного и доверенного ИИ внес АНО «Аналитический центр при Правительстве РФ». С 2021 года было поддержано 415 ИТ-проектов, 36 разработчиков открытых библиотек, 288 проектов для участия в акселерационной программе.

ФГАОУ ВО "Национальный исследовательский университет ИТМО" внес большой вклад в развитие сильного ИИ в России. Компания активно внедряет технологий сильного искусственного интеллекта (ИИ) за счет выполнения фундаментальных/прикладных исследований и разработок новых алгоритмов ИИ и создания на их основе программного обеспечения с элементами сильного ИИ для решения приоритетных задач цифровой трансформации и интеллектуализации промышленных производств в интересах индустриальных партнеров.

Институт системного программирования им. В.П. Иванникова Российской академии наук также активно ведет научную деятельность в области автоматического машинного обучения. Программа Центра нацелена на создание единой технической, методологической и организационной инфраструктуры для разработки доверенных систем, использующих технологии ИИ (ТИИ). Она включает в себя создание доверенных фреймворков машинного обучения, разработку принципиально новых программных инструментов для анализа, выявления и противодействия угрозам, которые специфичны для ТИИ, разработку инструментов для повышения интерпретируемости моделей машинного обучения и их результатов. Ряд технологий ИСП РАН уже включен в Реестр российского программного обеспечения:

- статический анализатор языка C# SharpChecker,
- статический анализатор Svace,
- система интеллектуального анализа текстов Texterra,
- цифровой испытательный стенд DigiTEF,
- облачная среда «Асперитас»,
- Talisman.Биография,
- Talisman.Поток,
- программная платформа организации проблемно-ориентированных Web-лабораторий FANLIGHT,
- инструмент определения поверхности атаки «Natch»,
- комплекс анализа Crusher,
- инструмент динамического анализа помеченных данных «Блесна».

Университет Иннополис — российская автономная некоммерческая организация высшего образования в городе Иннополисе (Верхнеуслонский район, Республика Татарстан), специализирующаяся на образовании, исследованиях и разработках в области информационных технологий и робототехники. Университет ведет передовые разработки в области нейротехнологий, такие как:

- Развитие физико-математических основ для диагностики состояния человека на основе данных электроэнцефалографии и ближнеинфракрасной спектроскопии,
- Физико-математические методы анализа функциональных связей головного мозга для диагностики моторных функций человека,
- Разработка методов анализа данных нейровизуализации для выявления процессов в головном мозге человека при моторной активности для развития интерфейсов мозг-компьютер,
- Персонализация траектории обучения на основе измерения когнитивных и личностных характеристик обучаемого с использованием нейроинтерфейса "мозг-компьютер" и методов искусственного интеллекта",
- Новые физико-математические методы, основанные на машинном обучении и рекуррентном анализе, для изучения нейронной активности человека при когнитивных процессах,
- Физические методы оценки и контроля когнитивных функций человека в процессе обучения.

ФГАОУ ВО "Национальный исследовательский университет – Высшая школа экономики" занимается межотраслевыми технологиями искусственного интеллекта и

искусственный интеллект для иных приоритетных отраслей экономики и социальной сферы.

Основными проектами Высшей школы экономики являются:

- Обучение, понимание и оптимизация моделей ИИ,
- ИИ в биоинформатике,
- Новые методы машинного обучения для оценки и интерпретации языковых моделей,
- Развитие методик и технологий прикладного ИИ,
- Мониторинг качества моделей в условиях сложности банковского регулирования и изменяющихся входных данных,
- Нейросетевые алгоритмы анализа динамики эмоционального состояния и вовлеченности учащихся на основе данных видеонаблюдения.

Центр компетенций НТИ МФТИ - это платформа, созданная в рамках национальной технологической инициативы (НТИ) и функционирующая на базе Московского физико-технического института (МФТИ). Его задача заключается в развитии инновационных проектов и технологий в области искусственного интеллекта, квантовых технологий, нанотехнологий, биотехнологий, космических технологий и других областей. Центр также обучает специалистов в этих областях и участвует в международных научных и технологически проектах. Основными проектами центра компетенций НТИ являются:

- Нейроморфные платформы: разработка архитектуры аппаратных нейроморфных платформ, аппаратная реализация искусственных нейронных сетей, создание нейронных сопроцессоров на основе мемристорных и оптогенетических технологий, создание прототипа нейроморфного вычислительного устройства,
- Сильный искусственный интеллект: разработка новых архитектур сильного машинного интеллекта на основе реверс-инжиниринга принципов работы мозга,
- Операционная среда роботов хоRDE: разработка архитектуры и макета, специализированной «операционной системы» для программирования роботов и робота-аватара.

Потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно-технической деятельности в Российской Федерации в части генеративного искусственного интеллекта являются промышленные, медицинские, транспортные и логистические, энергетические компании, а также компании занимающиеся рекламой и маркетингом, финансами и страхованием, медициной и биотехнологиями, производством ПО, также важно заметить заинтересованные государственные организации: министерство финансов, здравоохранения, образования и науки и др. Заинтересованными компаниями являются в технологиях AutoML: Тинькофф

банк, Газпром Нефть, МТС, Мегафон, АвтоВаз и др.

2.7 Системы обработки больших объёмов данных

Объём мирового рынка обработки больших данных в 2022 году оценивался в 271,83 млрд долларов США [260]. Прогнозируется, что рынок вырастет с 307,52 млрд долларов США в 2023 году до 745,15 млрд долларов США к 2030 году, демонстрируя среднегодовой рост в 13,5 % в течение прогнозируемого периода. Факторами роста рынка можно разделить на следующие категории:

— Увеличение объёма данных. Экспоненциальный рост данных, генерируемых различными источниками, такими как социальные сети, устройства интернет вещей и цифровые платформы, вызвал потребность в эффективных системах обработки больших данных;

— Растущее внедрение облачных сервисов. Внедрение облачных вычислений и масштабируемой инфраструктуры предприятиями для хранения и управления своими данными увеличило спрос на системы обработки больших данных.

— Спрос на аналитику в режиме реального времени. Потребность организаций в более быстром принятии решений на основе данных привела к спросу на системы обработки больших данных, которые позволяют проводить аналитику и анализ данных в реальном времени;

— Более низкая стоимость хранения данных. Падение цен на решения для хранения данных сделало организации более удобными для хранения и обработки больших объёмов данных;

— Правительственные инициативы. Правительства во всем мире принимают решения на основе данных и финансируют разработку технологий больших данных, стимулируя рост рынка систем обработки больших данных;

— Растущая индустрия электронной коммерции. Быстро развивающийся сектор электронной коммерции генерирует огромные объёмы транзакционных данных, которые необходимо обрабатывать и анализировать для обеспечения персонализированного обслуживания клиентов и оптимизации операций цепочки поставок;

— Растущие платформы социальных сетей. По мере того, как их использование продолжает расти, организациям нужны системы обработки больших данных для анализа огромных объёмов данных, созданных пользователями;

— Появление решений «большие данные как услуга». В настоящее время многие компании предлагают аналитику больших данных как услугу, что еще больше увеличивает доступность и спрос на системы обработки больших данных;

— Последние технологические достижения. Постоялая модернизация и разработка новых технологий, алгоритмов и платформ для более эффективной обработки больших объёмов данных.

По данным Международной корпорации обработки данных (IDC), к 2025 году 152 200 устройств Интернета вещей будут подключаться в минуту [260]. Возрастающие требования к подключенным устройствам ускоряют выполнение периферийных вычислений. Решения для периферийных вычислений определяются как структура, в которой процессоры расположены ближе к месту назначения или источнику данных, а не к облакам.

Доля рынка систем обработки больших данных по отраслям:

- аналитика клиентов – 13 %,
- аналитика цепочки поставок – 9 %,
- маркетинговая аналитика – 25 %,
- ценовая аналитика – 12 %,
- пространственный анализ – 10 %,
- анализ рисков и кредитоспособности – 22 %,
- транспортная аналитика – 4 %,
- другое – 5 %.

Список ключевых компаний, предоставляющие услуги в области систем обработки больших данных: IBM (США.), SAP SE (Германия), Microsoft (США), SAS Institute Inc. (США), Fair Isaac (США), Oracle (США), Salesforce Inc. (США), Equifax, Inc. (США), TransUnion (США), QlikTech International AB (США).

IBM – многонациональная технологическая компания, предлагающая ряд продуктов и услуг, связанных с системами обработки больших данных. Рассмотрим основные продукты компании в области больших данных:

— IBM InfoSphere DataStage – это мощный масштабируемый инструмент ETL, обеспечивающий интеграцию всех типов данных практически в реальном времени в локальных и облачных средах;

— Apache Hadoop – это программный проект с открытым исходным кодом, предназначенный для эффективной обработки больших пакетов данных;

— IBM InfoSphere BigInsights – этот продукт предоставляет предприятию всю мощь Hadoop и дополняет его производительностью, надежностью, безопасностью и административными функциями. Это помогает фирмам обнаруживать и анализировать новые бизнес-идеи, скрытые в больших наборах данных;

— IBM BigIntegrate – это гибкая и масштабируемая платформа для извлечения,

преобразования и интеграции данных Hadoop;

— IBM PureData System for Analytics – программное обеспечение и устройство для работы с данными, обеспечивающее высокопроизводительную аналитику для рабочих нагрузок больших данных;

— IBM Big Data and Analytics Platform – эта платформа помогает предприятиям собирать, интегрировать и анализировать информацию о клиентах. Он поддерживает обзор «шаблона клиента на 360°»;

— IBM Db2 Big SQL – программное обеспечение хранилища данных, которое поддерживает структурированные и неструктурированные данные и может обрабатывать большие объёмы данных;

— Apache Spark – это платформа параллельной обработки с открытым кодом, которая поддерживает обработку в памяти, чтобы повысить производительность приложений, анализирующих большие данные.

SAP SE – это мировой лидер в области бизнес-приложений и аналитики, основанная на обработке больших данных. Рассмотрим список основных продуктов SAP SE в области систем обработки больших данных:

— SAP IQ – это крупномасштабная система управления реляционными базами данных, используемая для обработки больших данных.

— SAP HANA – это система управления реляционными базами данных в памяти, ориентированная на столбцы;

— SAP Vora – механизм распределенных вычислений в памяти, разработанный специально для обработки и анализа больших данных;

— SAP Cloud Platform Big Data Services – облачная платформа, позволяющая пользователям обрабатывать и анализировать большие данные с помощью Apache Hadoop и Spark;

— SAP Predictive Analytics – программный инструмент, который позволяет пользователям создавать прогнозные модели и выполнять расширенный анализ больших наборов данных;

— SAP Data Intelligence – это облачная платформа управления данными, которая может обрабатывать несколько типов данных и предоставлять расширенную аналитику поверх данных. Платформа также использует машинное обучение для анализа данных.

Рассмотрим наиболее популярные продукты Microsoft в области систем обработки больших данных:

— Azure Data Factory – технология оркестровки, используемая для автоматизации рабочих процессов, которые преобразуют исходные данные, перемещают данные между

несколькими источниками и приемниками, загружают обработанные данные в хранилище аналитических данных или передают результаты прямо в отчет или панель мониторинга;

- Azure Analysis Services – модель данных централизованной онлайн-аналитической обработки, на которой могут основываться отчеты, информационные панели и интерактивный анализ;

- Power BI – сервис, позволяющий аналитикам данных создавать интерактивные визуализации данных на основе моделей данных;

- Microsoft Excel – программное обеспечение, которое предлагает множество возможностей анализа и визуализации данных. Аналитики данных могут использовать Excel для создания моделей данных документов из аналитических хранилищ данных или для извлечения данных из моделей данных OLAP в интерактивные сводные таблицы и диаграммы;

- Azure HDInsight – это полностью управляемая облачная служба, поддерживающая платформы обработки больших данных, такие как Hadoop и Spark, с возможностью масштабирования;

- Azure Stream Analytics – механизм обработки данных в реальном времени, который может принимать, обрабатывать и анализировать огромные объемы данных из различных источников;

- SQL Server – платформа данных, которая позволяет пользователям хранить, управлять и анализировать структурированные и неструктурированные данные.

Компания "SAS Institute" является одним из ведущих поставщиков программных решений и услуг в области анализа данных и бизнес-интеллекта. Ее продукты предназначены для работы с большими объемами данных и позволяют управлять, анализировать и использовать данные для принятия ключевых бизнес-решений. Рассмотрим основные продукты компании:

- SAS Viya – это облачная платформа для бизнес-аналитики, которая предлагает множество инструментов для обработки больших данных, включая машинное обучение, глубокое обучение и обработку естественного языка;

- SAS STAT – это программное обеспечение для статистических процессов с широким спектром функций;

- Event Stream Processing – это продукт, который позволяет проводить высокоскоростную обработку потоковых данных в реальном времени;

- High-Performance Analytics – это продукт, который позволяет проводить анализ больших данных с помощью параллельных вычислений;

- Data Quality – это продукт, который позволяет производить контроль качества данных и исправление ошибок в данных;

— Enterprise Miner – это продукт, который используется для построения, валидации и развертывания моделей данных;

— SAS Data Loader – этот продукт позволяет преобразовывать и очищать большие данные в кластерах Hadoop без необходимости написания кода.

— SAS Data Integration Studio – этот продукт представляет собой инструмент визуального проектирования, который позволяет создавать, внедрять и управлять процессами интеграции данных.

FICO – ведущая компания по разработке программного обеспечения для аналитики, которая использует системы больших данных и алгоритмы машинного обучения для предоставления прогнозной аналитики, для систем принятия решений и услуг кредитного скоринга. Основными продуктами компании являются:

— FICO Score – система оценки кредитоспособности, которая использует аналитику больших данных для оценки кредитного риска потребителей;

— FICO Decision Management Suite – платформа, позволяющая организациям автоматизировать процессы принятия решений с помощью аналитики больших данных;

— FICO Blaze Advisor – программное обеспечение для управления бизнес-правилами, которое использует аналитику больших данных, чтобы помочь организациям принимать более взвешенные решения;

— FICO Model Builder – программное обеспечение для моделирования данных, которое использует аналитику больших данных для создания прогнозных моделей для бизнеса;

— FICO Application Fraud Manager – программное обеспечение для обнаружения и предотвращения мошенничества, которое использует аналитику больших данных для выявления и пресечения мошеннических транзакций;

— FICO Debt Manager – программное обеспечение для сбора долгов, которое использует аналитику больших данных для оптимизации стратегий взыскания долгов;

— FICO Enterprise Fraud and Security Management – программное обеспечение для управления мошенничеством и безопасностью, которое использует аналитику больших данных для обнаружения и предотвращения мошенничества и нарушений безопасности.

Oracle – компания, которая предоставляет ряд аппаратных и программных решений для управления данными, включая обработку больших данных. Рассмотрим основные продукты компании:

— Oracle Big Data Cloud Service – облачный сервис для обработки и анализа больших объемов данных;

— Oracle Big Data Appliance – готовое решение для сбора, обработки и анализа

больших данных в реальном времени;

— Oracle NoSQL Database – высоко масштабируемая, распределенная база данных, предназначенная для хранения и обработки структурированных и неструктурированных данных;

— Oracle Event Hub Cloud Service – сервис обработки потока событий, который позволяет складывать входящие потоки событий в хранилище и обрабатывать их в реальном времени;

— Oracle Stream Analytics – инструмент для анализа потоков данных в реальном времени;

— Oracle Data Integrator for Big Data – инструмент для интеграции больших объемов данных, обеспечивающий загрузку, сбор, обработку и передачу данных между различными системами;

— Oracle DataRaker – программное обеспечение для анализа данных, используемое в промышленности для мониторинга и контроля за энергопотреблением;

— Oracle GoldenGate for Big Data – программное обеспечение для репликации и синхронизации данных между различными источниками данных;

— Oracle Big Data Spatial and Graph – программное обеспечение для анализа пространственных данных и графов;

— Oracle Business Intelligence Enterprise Edition – пакет программных продуктов для аналитики данных и бизнес-отчетности.

Российский рынок систем обработки больших объемов данных можно оценить в 200 млрд руб. на 2022 год. Как предсказывают аналитики к 2024 году объем рынка вырастет до 319 млрд. руб. [261]. Большую долю объема российского рынка занимают следующие компании: Яндекс, Ростелеком, VK, Сбербанк, Polymatica, INOSTUDIO, ISS Art, ZUZEX, Develonica, НПЦ «БизнесАвтоматика», Visiology, Modus BI, Webiomed, Metacommerce, Medframe и др.

Яндекс является лидирующей компанией на российском рынке систем обработки больших объемов данных. Рассмотрим основные продукты компании:

— Yandex Data Factory – эта платформа предлагает решения для машинного обучения и искусственного интеллекта для бизнеса, включая профилактическое обслуживание, прогнозирование спроса и обнаружение аномалий;

— Yandex ClickHouse – это система управления базами данных с открытым исходным кодом, которая может обрабатывать петабайты данных, сохраняя при этом высокую производительность запросов. Он предназначен для анализа больших наборов данных в режиме реального времени;

— Облако Яндекса – это платформа облачных вычислений, которая обеспечивает масштабируемую вычислительную мощность и хранилище для задач обработки данных. Она предлагает ряд услуг, таких как виртуальные машины, оркестрация контейнеров и хранилище объектов;

— Яндекс Метрика – это служба веб-аналитики, которая в режиме реального времени предоставляет данные о посещаемости веб-сайта, поведении пользователей и коэффициентах конверсии. Он предлагает ряд функций, таких как тепловые карты, воспроизведение сеанса и анализ воронки;

— Yandex DataLens – это инструмент визуализации данных, который позволяет пользователям создавать интерактивные информационные панели и отчеты. Он предлагает ряд вариантов визуализации, таких как диаграммы, карты и таблицы;

— Яндекс Контур – платформа управления данными, обеспечивающая обработку и анализ данных в режиме реального времени.

— Yandex Query Language – язык программирования, используемый для запросов и анализа больших наборов данных.

"Ростелеком" – это российская телекоммуникационная компания, которая занимается созданием и развитием инфраструктуры для передачи данных, как по традиционным сетям связи, так и через мобильные технологии. В рамках области систем обработки больших данных, "Ростелеком" предлагает решения для структурирования, хранения, анализа и использования больших объёмов информации, включая услуги по управлению данными, облачную аналитику и другие технологии. Эти услуги могут быть востребованы различными компаниями в разных отраслях для оптимизации процессов, повышения эффективности бизнеса и принятия решений на основе данных. Рассмотрим наиболее популярные продукты компании в области обработки больших объёмов данных:

— Rostelecom Analytics Platform – облачная платформа для обработки и анализа больших объёмов данных;

— Rostelecom Big Data Platform – комплексная инфраструктура для управления большими данными, включая хранение, обработку и анализ;

— Rostelecom Cloud Storage – безопасное и масштабируемое облачное хранилище для хранения и управления большими данными;

— Rostelecom Hadoop Solutions – набор инструментов и сервисов для построения и развертывания систем обработки больших данных на базе Hadoop;

VK – это российская инвестиционная технологическая корпорация, предоставляющая широкий спектр продуктов и услуг. Основными продуктами VK в области обработки больших объёмов данных являются:

— VK Cloud – это решение для хранения и обработки больших данных в облаке на основе Hadoop, Spark, ClickHouse;

— Mail.ru Machine Learning: рабочее место аналитика – сервис для быстрой разработки приложений на основе машинного обучения. Быстрое развёртывание сред обучения нейронных сетей. Поддержка matplotlib, TensorFlow, LightGBM (GBDT), PyTorch, XGBoost, pandas и других инструментов.

— PREDICT (Predictive Analytic Solutions) – проект VK, ключевой задачей которого является создание для клиентов продуктов и сервисов на основе предиктивных моделей, построенных с использованием методов машинного обучения. Прежде всего это проекты, призванные серьезно повысить эффективность процессов маркетинга и продаж, оптимизировать внутренние процессы крупных компаний, управление рисками, работу с персоналом и многое другое.

Сбербанк – это банк лидер в области систем обработки больших данных в России. Банк уже много лет разрабатывает и внедряет передовые технологии, связанные с аналитикой данных, например:

— SberData Platform – это платформа для управления большими данными, разработанная Сбербанком, которая предоставляет решения для сбора, хранения, обработки и анализа данных в больших объёмах;

— SberCloud – платформа облачных вычислений Сбербанка, предлагающая услуги по обработке больших данных на базе технологий Hadoop и Apache Spark;

— SberLog – система управления журналами, которая собирает, хранит и анализирует журналы в режиме реального времени. Платформа позволяет предприятиям активно отслеживать и устранять проблемы;

— SberStream – система потоковой аналитики в реальном времени, которая позволяет предприятиям обрабатывать и анализировать потоки данных в режиме реального времени. СберСтрим основан на Apache Kafka и Apache Flink;

— SberSQL – распределенный механизм запросов SQL, который позволяет предприятиям быстро и эффективно запрашивать большие объёмы данных. SberSQL основан на Apache Calcite;

— SberML – платформа машинного обучения, которая позволяет компаниям обучать, разворачивать модели машинного обучения и управлять ими в любом масштабе. SberML основан на TensorFlow и Apache Spark;

— SberSearch – поисковая система, которая позволяет компаниям искать и анализировать большие объёмы текстовых данных.

МТС является крупнейшим российским оператором связи, предоставляющим услуги

мобильной и фиксированной связи, выхода в интернет, телевидения, облачных услуг и других телекоммуникационных услуг. Из продуктов компании в области обработки больших данных можно выделить:

— CloudMTS – это платформа, предоставляющая облачные услуги для хранения и обработки данных, а также предоставляет решения для разработки и внедрения приложений в облаке;

— MTC Big Data Platform: это комплексная платформа больших данных, которая помогает организациям управлять большими объёмами структурированных и неструктурированных данных, обрабатывать их и извлекать ценную информацию. Он занимается сбором, обработкой, хранением и анализом данных, тем самым предоставляя универсальное решение для управления большими данными;

— MTC Data Warehouse: это масштабируемое решение для хранения данных, которое позволяет организациям хранить и управлять структурированными данными в петабайтном масштабе. Он обеспечивает встроенную поддержку инструментов и технологий на основе Hadoop, а также поддержку традиционных технологий хранения данных;

— MTC Data Lake: это облачное решение для озера данных, которое предоставляет единый репозиторий для всех типов данных, включая структурированные, частично структурированные и неструктурированные данные. Это позволяет организациям хранить, обрабатывать и анализировать большие объёмы данных без необходимости в дорогостоящей локальной инфраструктуре.

Polymatica – это компания–разработчик программного обеспечения, специализирующаяся на разработке и внедрении систем управления электронными документами и бизнес–процессами. Также компания имеет решения в области обработки больших данных, например, «Polymatica Analytics» программа для многомерного анализа показателей и метрик без потери детализации. Платформа основана на таких технологиях, как: машинное обучение, искусственный интеллект, анализ больших данных.

НПЦ «БизнесАвтоматика» – это производственный центр, специализирующийся на разработке и внедрении интеллектуальных систем автоматизации и управления бизнес–процессами. Главным продуктом в области обработки больших объёмов данных является Visary ETL. Visary ETL – надёжный и простой инструмент с графическим интерфейсом для управления пакетной обработкой данных и автоматизации потоковой передачи информации из одной системы в другую.

Visiology – российская ИТ–компания, разработчик программного обеспечения для анализа и визуализации данных. Основным продуктом компании является одноименный сервис «Visiology» для быстрого и эффективного решения задач сбора и обработки больших

объёмов данных. В апреле 2023 года компания представила обновление платформы Visiology 3 и Visiology 2, которые обеспечат пользователям более глубокий уровень разделения прав доступа и возможностей работы без авторизации. В свою очередь Visiology 3 дополнительно обладает достаточным функционалом для решения большей части типовых задач в сфере бизнес аналитики.

Modus BI – российский разработчик аналитических решений для корпоративного сектора и государственных органов власти. Основным продуктом компании является одноимённая платформа «Modus BI», которая способна визуализировать данные из различных источников, формировать отчетность и создавать прогнозы для эффективного управления всеми сферами деятельности.

Также стоит выделить университет Иннополис. Университет Иннополис — российская автономная некоммерческая организация высшего образования в городе Иннополисе (Верхнеуслонский район, Республика Татарстан), специализирующаяся на образовании, исследованиях и разработках в области информационных технологий и робототехники. На данный момент в Иннополисе ведется 8 проектов по блокчейну и облачным технологиям финансируемые грантами и индустриальными партнерами, например:

- Разработка и внедрение системы заправки воздушных судов, интегрированной с системой распределенного реестра с полной формальной верификацией,
- Разработка платформы для голосования на основе полностью верифицированной системы распределенного реестра Innochain Urban в сфере городского хозяйства,
- Цифровая модель города Иннополис,
- Разработка многофункциональной цифровой платформы по мониторингу КРС,
- Лидирующий исследовательский центр в области систем распределенного реестра.

Потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части систем обработки больших объёмов данных являются компании: пищевые, финансовые, медицинские, промышленные, энергетические, а также компании занимающиеся рекламой и маркетингом, электронной коммерцией, розничной торговлей, производством Интернет вещей также важно заметить заинтересованные государственные организации: министерство финансов, здравоохранения, образования и науки и др. Заинтересованным компания являются: Озон, Онлайн Трейд, Wildberries, МТС, Роснефть, РЖД и др.

3 АНАЛИЗ ИМЕЮЩИХСЯ РОССИЙСКИХ И ЗАРУБЕЖНЫХ РАЗРАБОТОК, КОТОРЫЕ ОКАЗЫВАЮТ (МОГУТ ОКАЗАТЬ) ВЛИЯНИЕ НА РАЗВИТИЕ СООТВЕТСТВУЮЩЕГО ПРИОРИТЕТА 20А, ДЛЯ ИХ СОПОСТАВЛЕНИЯ С ПРЕДЛОЖЕНИЯМИ, ПРЕДСТАВЛЯЕМЫМИ В АНАЛИЗИРУЕМЫХ ЗАЯВКАХ

Советом по приоритетному направлению 20а на заседании 22 июня 2023 г. были сформулированы предложения по формированию тематик заявок на разработку комплексных научно-технических проектов и программ, направляемых в Совет по приоритетному направлению 20а, которые оказывают существенное влияние на развитие новых рынков и применимы при разработке и реализации комплексных научно-технических проектов и программ по нескольким приоритетным направлениям:

- перспективные методы искусственного интеллекта, нейротехнологии;
- новые производственные технологии;
- технологии новых материалов и веществ;
- системы автоматизированного проектирования и управления жизненным циклом изделий;
- платформы управления технологическими и производственными процессами на базе технологий «интернета вещей» (IIoT) и распределенных реестров (DPP);
- технологии и компоненты робототехники и сенсорики;
- перспективные технологии беспроводной связи;
- квантовые вычисления и квантовые коммуникации;
- новое промышленное и общесистемное программное обеспечение;
- технологии виртуальной и дополненной реальности;
- анализ больших данных;
- вычислительная/суперкомпьютерная инфраструктура.

При этом необходимо отметить, что применение систем искусственного интеллекта прямо или косвенно является необходимым условием эффективной реализации комплексных научно-технических проектов и программ по указанным перспективным направлениям.

Кроме того, на заседании 22 июня 2023 г. было предложено оценить актуальность ранее рассмотренных КНТП и вынести на повторное рассмотрение наиболее значимые.

Ранее Советом были рассмотрены заявки на разработку КНТП по вышеуказанным перспективным направлениям:

По направлению «Технологии новых материалов и веществ»:

- проект КНТП МГУ им. М.В. Ломоносова с наименованием «Новые композитные материалы: технологии конструирования и производства» 11.05.2023 в

окончательной редакции был рассмотрен Советом.

Сопоставление проекта с имеющимися российскими и зарубежными разработками, о которых имеется в ограниченном открытом доступе из-за коммерческой тайны и закрытости разработок, показало, что он находится на мировом научно-техническом уровне [268], и проект КНТП был согласован для представления в Правительство РФ.

По направлению «Системы автоматизированного проектирования и управления жизненным циклом изделий»:

– заявка 2019-20(a)-9906-1009 АО «Алмаз-Антей» с наименованием «Разработка комплексной универсальной системы дистанционного мониторинга сложных технических систем и изделий в режиме реального времени» 26.02.2020 была рассмотрена Советом, и, при признании актуальности темы и использовании достижений искусственного интеллекта, инициатору было указано на недостаточную проработку прикладных применений результатов и имеющегося задела российских и зарубежных разработок, и было принято решение вернуть заявку на доработку. Доработанная заявка в Совет не поступала.

Необходимо отметить, что мировой рынок систем автоматизированного проектирования и управления жизненным циклом изделий достиг в 2022 г. 9 млрд долларов, российский рынок – 4 млрд рублей (60 млн долларов) [269, 270]. На 2023 год в России не хватает более 2000 отечественных САПР, что стало серьёзной проблемой для отечественной промышленности со времени введения санкций против РФ [271].

По направлению «Платформы управления технологическими и производственными процессами на базе технологий «интернета вещей» (IIoT) и распределенных реестров (TRP)»:

— заявка 2021-20(a)-1144-4164 ПАО «Межгосударственная акционерная компания «Вымпел» (ПАО «МАК «Вымпел») с наименованием «Национальная система дистанционного мониторинга, анализа статистики загрузки, структуры и состояния промышленного оборудования Российской Федерации». 17.11.2021 Совет рассмотрел заявку, признал её актуальность и практическую значимость, однако учитывая, что заявка частично соответствует требованиям, предъявляемым к комплексным научно-техническим проектам полного инновационного цикла рекомендовал инициатору доработать Заявку в соответствии с требованиями регламентирующих документов, а также с учётом высказанных в процессе обсуждения замечаний и предложений и повторно представить ее на очередном заседании Совета. Доработанная заявка в Совет не поступала;

— заявка 2020-20(a)-9587-7581 Государственной корпорации «Ростех» с наименованием «Комплексная разработка и производство приоритетных доверенных интеллектуальных программно-аппаратных платформ на основе отечественных электронных

компонентов и программного обеспечения» 07.05.2020 одобрена Советом с учётом её сопоставления с имеющимися российскими и зарубежными разработками и необходимости создания продукта, обеспечивающего технологическую независимость российских разработчиков в сложившейся политико-экономической ситуации. Инициатору было предложено сформировать предложение о разработке комплексной научно-технической программы и представить его на очередном заседании Совета. Предложение о разработке КНТП в Совет не поступало.

Необходимо отметить, что мировой рынок платформ управления технологическими и производственными процессами на базе технологий ИИТ и ТРР достиг в 2022 г. 88 млрд долларов, российский рынок – 148,5 млрд рублей (2 млрд долларов) [262, 263] и растёт на 18,3 % в год, что доказывает потребность РФ в данных технологиях [264].

По направлению «Технологии и компоненты робототехники и сенсорики»:

— предложение 2019-20(а)-4935-4784 ФГАНУ ЦНИИ РТК с наименованием «Робототехнические технологии вывода объектов атомной энергетики из эксплуатации» 21.10.2020 советом было одобрено доработанное предложение. Координационный совет рассмотрел предложение 22.05.2021, однако Комиссия по научно-технологическому развитию РФ от 07.02.2022 приостановила рассмотрение предложения до выхода новых нормативно-правовых актов, касающихся прохождения КНТП.

Следует отметить, что мировой рынок технологий и компонентов робототехники и сенсорики достиг в 2022 г. 50 млрд долларов в 2022 году, когда российский рынок достиг 1 млрд рублей (14 млн долларов) [272, 273]. В 2022 году российский рынок робототехники начал стремительно перестраиваться и, согласно последним исследованиям, плотность роботизации в России достигла 18 единиц, что соответствует среднегодовому росту в 165 % [274].

По направлению «Перспективные технологии беспроводной связи»:

— заявка 2020-20(а)-2664-1114 ФГАОУ ВО «Национальный исследовательский ядерный университет «МИФИ» с наименованием «Радиофотоника: перспективные системы и технологии для приёма, обработки и передачи информации» 07.05.2020 Совет одобрил представленную заявку, и инициатору было рекомендовано сформировать предложение о разработке КНТП с учётом замечаний, высказанных участниками заседания о:

необходимости формирования комплексного плана научных исследований, предусматривающего проведение междисциплинарных фундаментальных исследований;

расширении состава участников программы за счёт привлечения к её выполнению организаций, имеющих признанный научно-технический задел;

необходимости создания специализированных САПР, ориентированных на разработку отечественной радиодатонной компонентной базы, и её последующего создания.

Сформированное предложение в Совет не поступало.

Следует отметить, что мировой рынок перспективных технологий беспроводной связи достиг в 2022 г. 109 млрд долларов, когда российский рынок достиг 1,2 млрд долларов [265, 266]. Как показывают последние исследования, внедрение новых технологий беспроводной связи может охватить до 80 % российского рынка [267].

Анализ возможных причин задержки дальнейшего продвижения инициаторами многих заявок на разработку КНТП показывает, что они вызваны существенным снижением экономической активности из-за ситуацией с COVID-19 в 2020 – 2021 гг., наложением на Россию санкций недружественными государствами, концентрацией с 2022 г. ресурсов промышленности на заказах, связанных с СВО, а также ожиданием изменения нормативно-правовой базы, декларированных Комиссией по научно-технологическому развитию РФ от 7 февраля 2022 г. и уточнения Комиссией от 28 февраля 2023 г. срока их представления – 1 августа 2023 г.

ЗАКЛЮЧЕНИЕ

В настоящем аналитическом докладе были выявлены основные тенденции развития сферы науки и технологий по направлению 20а:

— основными тенденциями развития искусственного интеллекта в компьютерном зрении являются внедрения и разработка: детекции и идентификации объектов для беспилотного транспорта, детекция и идентификация объектов в промышленности, детекция и идентификация объектов в здравоохранении, компьютерное зрение в хирургии, в диагностировании патологий по изображению, компьютерное зрение в мобильных роботах, в промышленных роботах;

— основными тенденциями развития искусственного интеллекта в обработке естественного языка являются внедрения и разработка: генерация естественного языка в здравоохранении, статистическая обработка естественного языка, гибридная обработка естественного языка, обработка естественного языка на основе правил;

— основными тенденциями развития искусственного интеллекта в распознавании и синтезе речи являются глубокое обучение и нейронные сети, многоязычная и многодialeктная поддержка, интеграция с другими технологиями и сервисами в распознавании речи, пользовательский голос и персонализация в синтезе речи;

— основными тенденциями развития искусственного интеллекта в рекомендательных системах и интеллектуальных системах поддержки принятия решений являются объяснимость рекомендательных систем, гибридные рекомендательные системы, интеграция с Интернет вещей, планирование и принятие решений в промышленности, прогнозная аналитика;

— основными тенденциями развития генеративного искусственного интеллекта являются генерация контента в компьютерной графике и дизайне, генерация отчетов и документации в медицине, медицинская диагностика с использованием генеративно-состязательных сетях, обработка и генерация больших объёмов данных в здравоохранении, оптимизация и ускорение процессов в промышленности, оптимизация в розничной торговле с использованием потоковых моделей, чат-боты в маркетинге и рекламе, обработка и генерация больших объёмов данных в здравоохранении, анализ и прогнозирование в финансах и банковском деле в моделях с использованием моделей на базе трансформера;

— основными тенденциями развития автоматического машинного обучения являются глубокое обучение, генетические алгоритмы, автоэнкодеры в автоматизированной генерации признаков, решетчатый поиск, случайный поиск, оптимизация методов последовательного моделирования в оптимизации гиперпараметров, обучение с подкреплением, робототехника, предсказание финансовых рядов с использованием

нейронных сетей, нейроуправление и сильный искусственный интеллект.

— основными тенденциями развития систем обработки больших объёмов данных являются здравоохранение, промышленность в обработке больших объёмов данных в режиме реального времени, облачная обработка больших объёмов данных в здравоохранение и медико–биологических науках, розничной торговле и электронной коммерции, промышленной Интернет вещей, инструменты машинного обучения для интеграции и управления большими данными, распределенные вычисления и обработка данных в системах обработки больших объёмов данных с использованием машинного обучения;

— рынок технического зрения будет расти и расширяться как в мире, так и в России. В работе были выявлены основные ключевые российские и зарубежные разработчики технического зрения. В настоящее время в России существуют готовые отечественные решения, которые могут решить широкий спектр задач начиная от промышленности заканчивая розничной торговлей. Был оценён объём российского и зарубежного рынка, а также определены перспективы роста более чем на 5 лет;

— рынок обработки естественного языка будет расти и расширяться как в мире, так и в России. В работе были выявлены основные ключевые российские и зарубежные разработчики технического зрения. В настоящее время в России существуют готовые отечественные решения, которые могут решить широкий спектр задач начиная от промышленности заканчивая розничной торговлей. Был оценён объём российского и зарубежного рынка, а также определены перспективы роста более чем на 5 лет;

— потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части искусственного интеллекта распознавании и синтезе речи являются государственные органы и ведомства, исследовательские и научно–образовательные учреждения, разработчики робототехники и автономных систем, программного обеспечения в области машинного обучения, крупные производственные предприятия, крупные банки и электронная и розничная торговля. Из заинтересованных компаний можно выделить Газпром-Нефть, X5 Group, Северсталь, ПИК, МТС и др.;

— потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части искусственного интеллекта в рекомендательных системах и интеллектуальных системах поддержки принятия решений являются государственные органы и ведомства, исследовательские и научно–образовательные учреждения, разработчики робототехники и автономных систем, программного обеспечения в области машинного обучения, крупные производственные предприятия, крупные банки и электронная и

розничная торговля. Из заинтересованных компаний можно выделить X5 Group, Роснефть, ПИК, АвтоВаз, КАМАЗ;

— потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части генеративного искусственного интеллекта являются компании: промышленные, медицинские, транспортные, энергетические, а также компании занимающиеся рекламой и маркетингом, дизайном и архитектурой, производством ПО, также важно заметить заинтересованные государственные организации: министерство финансов, здравоохранения, образования и науки и др. Заинтересованным компания являются: Сбербанк, VK, Озон, Медиаскоп, Тинькофф, X5 GROUP, Мегафон и др.;

— потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части генеративного искусственного интеллекта являются компании: промышленные, медицинские, транспортные и логистические, энергетические, а также компании занимающиеся рекламой и маркетингом, финансами и страхованием, медициной и биотехнологиями, производством ПО, также важно заметить заинтересованные государственные организации: министерство финансов, здравоохранения, образования и науки и др. Заинтересованными компаниями являются в технологиях AutoML: Тинькофф банк, Газпром-Нефть, МТС, Мегафон, АвтоВАЗ и др.;

— потенциальными заказчиками комплексных программ и комплексных проектов с использованием результатов научной и (или) научно–технической деятельности в Российской Федерации в части систем обработки больших объёмов данных являются пищевые, финансовые, медицинские, промышленные, энергетические компании, а также компании, занимающиеся рекламой и маркетингом, электронной коммерцией, розничной торговлей, производством Интернет вещей Также важно отметить заинтересованные государственные организации: министерства финансов, здравоохранения, науки и высшего образования и др.

Сопоставление имеющихся российских и зарубежных разработок, которые могут оказать влияние на развитие соответствующего приоритета 20а, проведённое Советом в процессе рассмотрения поданных в Совет заявок и предложений на разработку КНТП, показал, что предлагаемые технические решения в ряде случаев требуют дополнительного учёта перспективных направлений и новых научно-технических и(или) технологических решений и условий, в связи с чем Совет направил их на доработку.

СПИСОК ИСТОЧНИКОВ

1. Google Scholar [Электронный ресурс]. URL: <https://scholar.google.ru/schhp?hl=en> (дата обращения: 18.05.2023).
2. Publications - Dimensions [Электронный ресурс]. URL: <https://app.dimensions.ai/discover/publication> (дата обращения: 18.05.2023).
3. Ali M.H. и др. Autonomous vehicles decision-making enhancement using self-determination theory and mixed-precision neural networks // *Multimed. Tools Appl.* Springer, 2023. С. 1–24.
4. Alfred Daniel J. и др. Fully convolutional neural networks for LIDAR–camera fusion for pedestrian detection in autonomous vehicle // *Multimed. Tools Appl.* Springer, 2023. С. 1–24.
5. Qiao Y., Yin J., Wang W. Survey of Deep Learning for Autonomous Surface Vehicles in Marine Environments // *ieeexplore.ieee.org*. 2023.
6. Sharma T. и др. Deep Learning-Based Object Detection and Scene Perception under Bad Weather Conditions // *Electron. MDPI*, 2022. Т. 11, № 4. С. 563.
7. Apostolopoulos I.D., Tzani M.A. Industrial object and defect recognition utilizing multilevel feature extraction from industrial scenes with Deep Learning approach // *J. Ambient Intell. Humaniz. Comput.* Springer Science and Business Media Deutschland GmbH, 2022. Т. 1. С. 1–14.
8. Ullah W. и др. Artificial Intelligence of Things-assisted two-stream neural network for anomaly detection in surveillance Big Video Data // *Futur. Gener. Comput. Syst.* North-Holland, 2022. Т. 129. С. 286–297.
9. Chiu M.-C., Tsai H.-Y., Chiu J.-E. A novel directional object detection method for piled objects using a hybrid region-based convolutional neural network // *Adv. Eng. Informatics.* Elsevier, 2022. Т. 51. С. 101448.
10. Ashafuddula N.I.M., Islam R. Melanoma Skin Cancer and Nevus Mole Classification using Intensity Value Estimation with Convolutional Neural Network. 2022.
11. Zhao Y. и др. Identification of gastric cancer with convolutional neural networks: a systematic review // *Multimed. Tools Appl.* Springer, 2022. Т. 81, № 8. С. 11717–11736.
12. Warin K. и др. AI-based analysis of oral lesions using novel deep convolutional neural networks for early detection of oral cancer // *PLoS One. Public Library of Science*, 2022. Т. 17, № 8. С. e0273508.
13. Tyagi N. и др. Skin Cancer Prediction using Machine Learning and Neural Networks // *Proc. 5th Int. Conf. Contemp. Comput. Informatics, IC3I 2022.* Institute of Electrical and Electronics Engineers Inc., 2022. С. 271–275.
14. Zareian F., Rezaei N. Application of Artificial Intelligence in Lung Cancer Detection: The

- Integration of Computational Power and Clinical Decision-Making. Springer, Cham, 2022. С. 1–14.
15. Patel B., Makaryus A.N. Artificial Intelligence Advances in the World of Cardiovascular Imaging // *Healthc.* 2022, Vol. 10, Page 154. Multidisciplinary Digital Publishing Institute, 2022. Т. 10, № 1. С. 154.
 16. Zhang J. и др. Artificial Intelligence in Cardiovascular Atherosclerosis Imaging // *J. Pers. Med.* 2022, Vol. 12, Page 420. Multidisciplinary Digital Publishing Institute, 2022. Т. 12, № 3. С. 420.
 17. Ma P., Li Q., Li J. Application of Artificial Intelligence in Cardiovascular Imaging // *J. Healthc. Eng.* Hindawi Limited, 2022. Т. 2022.
 18. Huyut M.T., Velichko A. Diagnosis and Prognosis of COVID-19 Disease Using Routine Blood Values and LogNNet Neural Network // *Sensors* 2022, Vol. 22, Page 4820. Multidisciplinary Digital Publishing Institute, 2022. Т. 22, № 13. С. 4820.
 19. Kathamuthu N.D. и др. A deep transfer learning-based convolution neural network model for COVID-19 detection using computed tomography scan images for medical applications // *Adv. Eng. Softw.* Elsevier Ltd, 2023. Т. 175.
 20. Shanbehzadeh M., Nopour R., Kazemi-Arpanahi H. Developing an artificial neural network for detecting COVID-19 disease // *J. Educ. Health Promot.* Wolters Kluwer -- Medknow Publications, 2022. Т. 11, № 1. С. 2.
 21. Hu Q. и др. Explainable artificial intelligence-based edge fuzzy images for COVID-19 detection and identification // *Appl. Soft Comput.* Elsevier, 2022. Т. 123. С. 108966.
 22. Chen Y. и др. A Neural Network-Based Navigation Approach for Autonomous Mobile Robot Systems // *Appl. Sci.* 2022, Vol. 12, Page 7796. Multidisciplinary Digital Publishing Institute, 2022. Т. 12, № 15. С. 7796.
 23. Rubtsova J. Approach to Image-Based Segmentation of Complex Surfaces Using Machine Learning Tools During Motion of Mobile Robots // *Smart Innov. Syst. Technol.* Springer Science and Business Media Deutschland GmbH, 2022. Т. 232. С. 191–200.
 24. Sleaman W.K., Hameed A.A., Jamil A. Monocular vision with deep neural networks for autonomous mobile robots navigation // *Optik (Stuttg).* Urban & Fischer, 2023. Т. 272. С. 170162.
 25. Gao T. и др. An operational calibration approach of industrial robots through a motion capture system and an artificial neural network ELM // *Int. J. Adv. Manuf. Technol.* Springer Science and Business Media Deutschland GmbH, 2023. Т. 125, № 11. С. 5135–5147.
 26. Wang Y. и др. Automatic Training Method of Deep Neural Network for Robot Vision // *Lect. Notes Electr. Eng.* Springer Science and Business Media Deutschland GmbH, 2022. Т. 805

- LNEE. С. 568–579.
27. Yun J. и др. Grasping Pose Detection for Loose Stacked Object Based on Convolutional Neural Network with Multiple Self-Powered Sensors Information // *IEEE Sens. J.* Institute of Electrical and Electronics Engineers Inc., 2022.
 28. Bilal D.K. и др. Development of a vision based pose estimation system for robotic machining and improving its accuracy using LSTM neural networks and sparse regression // *Robot. Comput. Integr. Manuf.* Pergamon, 2022. Т. 74. С. 102262.
 29. Zheng C. и др. Hybrid offline programming method for robotic welding systems // *Robot. Comput. Integr. Manuf.* Pergamon, 2022. Т. 73. С. 102238.
 30. Horvath D. и др. Object Detection Using Sim2Real Domain Randomization for Robotic Applications // *IEEE Trans. Robot.* Institute of Electrical and Electronics Engineers Inc., 2022.
 31. Soffer S. и др. Artificial Intelligence for Interstitial Lung Disease Analysis on Chest Computed Tomography: A Systematic Review // *Acad. Radiol.* Elsevier, 2022. Т. 29. С. S226–S235.
 32. Liu H. и др. Artificial Intelligence-Based Breast Cancer Diagnosis Using Ultrasound Images and Grid-Based Deep Feature Generator // *Int. J. Gen. Med.* Dove Medical Press Ltd, 2022. Т. 15. С. 2271–2282.
 33. Nawaz S.A., Khan D.M., Qadri S. Brain Tumor Classification Based on Hybrid Optimized Multi-features Analysis Using Magnetic Resonance Imaging Dataset // <https://doi.org/10.1080/08839514.2022.2031824>. Taylor & Francis, 2022. Т. 36, № 1.
 34. Puneet, Kumar R., Gupta M. Optical coherence tomography image based eye disease detection using deep convolutional neural network // *Heal. Inf. Sci. Syst.* Springer, 2022. Т. 10, № 1. С. 1–16.
 35. Wu L. и др. Real-time artificial intelligence for detecting focal lesions and diagnosing neoplasms of the stomach by white-light endoscopy (with videos) // *Gastrointest. Endosc.* Mosby, 2022. Т. 95, № 2. С. 269-280.e6.
 36. Mascagni P. и др. Artificial Intelligence for Surgical Safety : Automatic Assessment of the Critical View of Safety in Laparoscopic Cholecystectomy Using Deep Learning // *Ann. Surg.* Wolters Kluwer, 2021.
 37. Ramesh S. и др. Dissecting Self-Supervised Learning Methods for Surgical Computer Vision. 2022. С. 1.
 38. Xu H. и др. Detection of blood stains using computer vision-based algorithms and their association with postoperative outcomes in thoracoscopic lobectomies // *Eur. J. Cardio-Thoracic Surg.* Oxford Academic, 2022. Т. 62, № 5. С. 154.
 39. Yuan H. и др. BioBART: Pretraining and Evaluation of A Biomedical Generative Language Model // *Proc. Annu. Meet. Assoc. Comput. Linguist.* Association for Computational

- Linguistics (ACL), 2022. С. 97–109.
40. Velichko A. и др. Complex Paralinguistic Analysis of Speech: Predicting Gender, Emotions and Deception in a Hierarchical Framework // INTERSPEECH 2022. International Speech Communication Association, 2022. Т. 2022-September. С. 4735–4739.
 41. August T. и др. Paper Plain: Making Medical Research Papers Approachable to Healthcare Consumers with Natural Language Processing // arxiv.org. 2022.
 42. Li M. и др. Auxiliary signal-guided knowledge encoder-decoder for medical report generation // World Wide Web. Springer, 2023. Т. 26, № 1. С. 253–270.
 43. Palaniyappan L. и др. Studying psychosis using Natural Language Generation: A review of emerging opportunities // Biol. Psychiatry Cogn. Neurosci. Neuroimaging. Elsevier, 2023.
 44. Yang K., Deng J., Chen D. Generating Natural Language Proofs with Verifier-Guided Search // Proc. 2022 Conf. Empir. Methods Nat. Lang. Process. EMNLP 2022. Association for Computational Linguistics (ACL), 2022. С. 89–105.
 45. Yan S., Zhu T. Disentangled Variational Topic Inference for Topic-Accurate Financial Report Generation. 2022. С. 18–24.
 46. Taha A., Cosgrave B., McKeever S. Using Feature Selection with Machine Learning for Generation of Insurance Insights // Appl. Sci. 2022, Vol. 12, Page 3209. Multidisciplinary Digital Publishing Institute, 2022. Т. 12, № 6. С. 3209.
 47. Reisenbichler M. и др. Frontiers: Supporting Content Marketing with Natural Language Generation // <https://doi.org/10.1287/mksc.2022.1354>. INFORMS, 2022. Т. 41, № 3. С. 441–452.
 48. Zhang H. и др. A Survey of Controllable Text Generation using Transformer-based Pre-trained Language Models. 2022.
 49. Zhang X. и др. Automatic Product Copywriting for E-commerce // Proc. AAAI Conf. Artif. Intell. Association for the Advancement of Artificial Intelligence, 2022. Т. 36, № 11. С. 12423–12431.
 50. Shankar V., Parsana S. An overview and empirical comparison of natural language processing (NLP) models and an introduction to and empirical application of autoencoder models in marketing // J. Acad. Mark. Sci. Springer, 2022. Т. 50, № 6. С. 1324–1350.
 51. Ganguly I. и др. Automated Error Labeling in Radiation Oncology via Statistical Natural Language Processing // Diagnostics 2023, Vol. 13, Page 1215. Multidisciplinary Digital Publishing Institute, 2023. Т. 13, № 7. С. 1215.
 52. Nijhawan T., Attigeri G., Ananthakrishna T. Stress detection using natural language processing and machine learning over social interactions // J. Big Data. Springer Science and Business Media Deutschland GmbH, 2022. Т. 9, № 1. С. 1–24.

53. Yaeger J.P. и др. Derivation of a natural language processing algorithm to identify febrile infants // *J. Hosp. Med.* John Wiley & Sons, Ltd, 2022. Т. 17, № 1. С. 11–18.
54. Gunter D. и др. Rule-based natural language processing for automation of stroke data extraction: a validation study // *Neuroradiology.* Springer Science and Business Media Deutschland GmbH, 2022. Т. 64, № 12. С. 2357–2362.
55. Yang Z. и др. Assessment of Natural Language Processing Methods for Ascertaining the Expanded Disability Status Scale Score From the Electronic Health Records of Patients With Multiple Sclerosis: Algorithm Development and Validation Study // *JMIR Med Inf.* 2022;10(1)e25157 <https://medinform.jmir.org/2022/1/e25157>. *JMIR Medical Informatics*, 2022. Т. 10, № 1. С. e25157.
56. Anbukkarasi S., Varadhaganapathy S. Neural network-based error handler in natural language processing // *Neural Comput. Appl.* Springer Science and Business Media Deutschland GmbH, 2022. Т. 34, № 23. С. 20629–20638.
57. Lai K. и др. A Natural Language Processing Approach to Understanding Context in the Extraction and GeoCoding of Historical Floods, Storms, and Adaptation Measures // *Inf. Process. Manag.* Pergamon, 2022. Т. 59, № 1. С. 102735.
58. Samant R.M. и др. Framework for Deep Learning-Based Language Models Using Multi-Task Learning in Natural Language Understanding: A Systematic Literature Review and Future Directions // *IEEE Access.* Institute of Electrical and Electronics Engineers Inc., 2022. Т. 10. С. 17078–17097.
59. Fogli D. и др. A hybrid approach to user-oriented programming of collaborative robots // *Robot. Comput. Integr. Manuf.* Pergamon, 2022. Т. 73. С. 102234.
60. Kaur G., Sharma A. A deep learning-based model using hybrid feature extraction approach for consumer sentiment analysis // *J. Big Data.* Springer Science and Business Media Deutschland GmbH, 2023. Т. 10, № 1. С. 1–23.
61. Dua S. и др. Developing a Speech Recognition System for Recognizing Tonal Speech Signals Using a Convolutional Neural Network // *Appl. Sci.* 2022, Vol. 12, Page 6223. *Multidisciplinary Digital Publishing Institute*, 2022. Т. 12, № 12. С. 6223.
62. Oruh J., Viriri S., Adegun A. Long Short-Term Memory Recurrent Neural Network for Automatic Speech Recognition // *IEEE Access.* Institute of Electrical and Electronics Engineers Inc., 2022. Т. 10. С. 30069–30079.
63. Trinh Van L. и др. Emotional Speech Recognition Using Deep Neural Networks // *Sensors* 2022, Vol. 22, Page 1414. *Multidisciplinary Digital Publishing Institute*, 2022. Т. 22, № 4. С. 1414.
64. Zhou L. и др. A CONFIGURABLE MULTILINGUAL MODEL IS ALL YOU NEED TO

- RECOGNIZE ALL LANGUAGES // ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. Institute of Electrical and Electronics Engineers Inc., 2022. T. 2022-May. C. 6422–6426.
65. Hussein A. и др. Textual Data Augmentation for Arabic-English Code-Switching Speech Recognition // 2022 IEEE Spok. Lang. Technol. Work. SLT 2022 - Proc. Institute of Electrical and Electronics Engineers Inc., 2023. C. 777–784.
 66. Dan Z. и др. Multi-task Learning with Auxiliary Cross-attention Transformer for Low-Resource Multi-dialect Speech Recognition // Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). Springer Science and Business Media Deutschland GmbH, 2022. T. 13551 LNAI. C. 107–118.
 67. Shashidhar R., Patilkulkarni S., Puneeth S.B. Combining audio and visual speech recognition using LSTM and deep convolutional neural network // Int. J. Inf. Technol. Springer Science and Business Media B.V., 2022. T. 14, № 7. C. 3425–3436.
 68. Jayanthi K., Mohan S., B L. An integrated framework for emotion recognition using speech and static images with deep classifier fusion approach // Int. J. Inf. Technol. Springer Science and Business Media B.V., 2022. T. 14, № 7. C. 3401–3411.
 69. Chen H., Chen S., Zhao J. Integrated Design of Financial Self-Service Terminal Based on Artificial Intelligence Voice Interaction // Front. Psychol. Frontiers Media S.A., 2022. T. 13.
 70. Rogowski A., Bieliszczuk K., Rapcewicz J. Integration of Industrially-Oriented Human-Robot Speech Communication and Vision-Based Object Recognition // Sensors (Switzerland). MDPI AG, 2020. T. 20, № 24. C. 1–22.
 71. Cornille T., Wang F., Bekker J. INTERACTIVE MULTI-LEVEL PROSODY CONTROL FOR EXPRESSIVE SPEECH SYNTHESIS // ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. Institute of Electrical and Electronics Engineers Inc., 2022. T. 2022-May. C. 8312–8316.
 72. Tu J. и др. Contextual Expressive Text-to-Speech // arxiv.org. 2022.
 73. James J. и др. Exploring Prosodic Features Modelling for Secondary Emotions Needed for Empathetic Speech Synthesis // Sensors. MDPI AG, 2023. T. 23, № 6. C. 2999.
 74. Matsunaga Y. и др. Empirical Study Incorporating Linguistic Knowledge on Filled Pauses for Personalized Spontaneous Speech Synthesis // Proc. 2022 Asia-Pacific Signal Inf. Process. Assoc. Annu. Summit Conf. APSIPA ASC 2022. Institute of Electrical and Electronics Engineers Inc., 2022. C. 1898–1903.
 75. Wang J. и др. RESIDUAL-GUIDED PERSONALIZED SPEECH SYNTHESIS BASED ON FACE IMAGE // ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. Institute of Electrical and Electronics Engineers Inc., 2022. T. 2022-May. C. 4743–4747.

76. Matsunaga Y. и др. Spontaneous speech synthesis with linguistic-speech consistency training using pseudo-filled pauses // arxiv.org. 2022.
77. Wang G. и др. An Audio-Visual Attention Based Multimodal Network for Fake Talking Face Videos Detection // arxiv.org. 2022.
78. Chen S. и др. Talking Head Generation Driven by Speech-Related Facial Action Units and Audio- Based on Multimodal Representation Fusion // arxiv.org. 2022.
79. Ashok K. и др. Collaborative analysis of audio-visual speech synthesis with sensor measurements for regulating human–robot interaction // Int. J. Syst. Assur. Eng. Manag. Springer, 2022. Т. 1. С. 1–8.
80. Markchom T., Liang H., Ferryman J. Scalable and explainable visually-aware recommender systems // Knowledge-Based Syst. Elsevier, 2023. Т. 263. С. 110258.
81. Alhejaili A., Fatima S. Expressive Latent Feature Modelling for Explainable Matrix Factorisation-based Recommender Systems // ACM Trans. Interact. Intell. Syst. ACM-PUB27New York, NY, 2022. Т. 12, № 3.
82. Balloccu G. и др. Post Processing Recommender Systems with Knowledge Graphs for Recency, Popularity, and Diversity of Explanations // SIGIR 2022 - Proc. 45th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. Association for Computing Machinery, Inc, 2022. С. 646–656.
83. Walek B., Fajmon P. A hybrid recommender system for an online store using a fuzzy expert system // Expert Syst. Appl. Pergamon, 2023. Т. 212. С. 118565.
84. Tanwar A., Vishwakarma D.K. A deep neural network-based hybrid recommender system with user-user networks // Multimed. Tools Appl. Springer, 2022. Т. 82, № 10. С. 15613–15633.
85. Somya R., Winarko E., Priyanta S. A hybrid recommender system based on customer behavior and transaction data using generalized sequential pattern algorithm // Bull. Electr. Eng. Informatics. Institute of Advanced Engineering and Science, 2022. Т. 11, № 6. С. 3422–3432.
86. Bouazza H., Said B., Zohra Laallam F. A hybrid IoT services recommender system using social IoT // J. King Saud Univ. - Comput. Inf. Sci. Elsevier, 2022. Т. 34, № 8. С. 5633–5645.
87. Ahmed M. и др. Rating-Based Recommender System Based on Textual Reviews Using IoT Smart Devices // Mob. Inf. Syst. Hindawi Limited, 2022. Т. 2022.
88. Beshley M. и др. Data Optimization for Industrial IoT-Based Recommendation Systems // Electron. 2023, Vol. 12, Page 33. Multidisciplinary Digital Publishing Institute, 2022. Т. 12, № 1. С. 33.
89. Deloitte. Trustworthy Artificial Intelligence [Электронный ресурс]. 2022. URL: <https://www2.deloitte.com/us/en/pages/deloitte-analytics/solutions/ethics-of-ai->

- framework.html (дата обращения: 26.06.2023).
90. Molodetskikh I. и др. Temporally coherent person matting trained on fake-motion dataset // *Digit. Signal Process. Academic Press*, 2022. Т. 126. С. 103521.
 91. Ivanova A. и др. Oracle Complexity Separation in Convex Optimization // *J. Optim. Theory Appl. Springer*, 2022. Т. 193, № 1–3. С. 462–490.
 92. Novitskii V., Gasnikov A. Improved exploitation of higher order smoothness in derivative-free optimization // *Optim. Lett. Springer Science and Business Media Deutschland GmbH*, 2022. Т. 16, № 7. С. 2059–2071.
 93. Makhalova T. и др. Introducing the closure structure and the GDPM algorithm for mining and understanding a tabular dataset // *Int. J. Approx. Reason. Elsevier*, 2022. Т. 145. С. 75–90.
 94. Krivonosov M. и др. Age-related trajectories of DNA methylation network markers: A parenclitic network approach to a family-based cohort of patients with Down Syndrome // *Chaos, Solitons & Fractals. Pergamon*, 2022. Т. 165. С. 112863.
 95. Krivonosov M.I. и др. A new cognitive clock matching phenotypic and epigenetic ages // *Transl. Psychiatry 2022 121. Nature Publishing Group*, 2022. Т. 12, № 1. С. 1–9.
 96. Vazhentsev A. и др. Uncertainty Estimation of Transformer Predictions for Misclassification Detection // *Proc. Annu. Meet. Assoc. Comput. Linguist. Association for Computational Linguistics (ACL)*, 2022. Т. 1. С. 8237–8252.
 97. Vasilkovsky M. и др. DetIE: Multilingual Open Information Extraction Inspired by Object Detection // *Proc. AAAI Conf. Artif. Intell. Association for the Advancement of Artificial Intelligence*, 2022. Т. 36, № 10. С. 11412–11420.
 98. Antsiferova A. и др. Video compression dataset and benchmark of learning-based video-quality metrics // *Adv. Neural Inf. Process. Syst. 2022*. Т. 35. С. 13814–13825.
 99. Kovalev KAUST D. и др. Accelerated Primal-Dual Gradient Method for Smooth and Convex-Concave Saddle-Point Problems with Bilinear Coupling // *Adv. Neural Inf. Process. Syst. 2022*. Т. 35. С. 21725–21737.
 100. Sadiev A., Kovalev D., Richtárik P. Communication Acceleration of Local Gradient Methods via an Accelerated Primal-Dual Algorithm with Inexact Prox // *arxiv.org*. 2022.
 101. Maciej Serda и др. A Damped Newton Method Achieves Global $O(1/k^2)$ and Local Quadratic Convergence Rate // *Uniw. śląski / под ред. G. Balint и др. arXiv*, 2022. Т. 7, № 1. С. 343–354.
 102. Tsvigun A. и др. Towards Computationally Feasible Deep Active Learning // *Find. Assoc. Comput. Linguist. NAACL 2022 - Find. Association for Computational Linguistics (ACL)*, 2022. С. 1198–1218.
 103. Tsvigun A. и др. Active Learning for Abstractive Text Summarization // *Find. Assoc. Comput.*

- Linguist. EMNLP 2022. Association for Computational Linguistics (ACL), 2023. С. 5157–5181.
104. Tian Y. и др. Acceleration in Distributed Optimization under Similarity. PMLR, 2022. С. 5721–5756.
 105. Liu J. и др. Multi-ship collision avoidance decision-making and coordination mechanism in Mixed Navigation Scenarios // Ocean Eng. Pergamon, 2022. Т. 257. С. 111666.
 106. Carvalho H.S. и др. An Intelligent Decision Support System for Road Freight Transport // Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics). Springer Science and Business Media Deutschland GmbH, 2022. Т. 13756 LNCS. С. 146–156.
 107. Samadi Gharajeh M., Jond H.B. An intelligent approach for autonomous mobile robots path planning based on adaptive neuro-fuzzy inference system // Ain Shams Eng. J. Elsevier, 2022. Т. 13, № 1. С. 101491.
 108. Li J., Zhou G., Zhang C. A twin data and knowledge-driven intelligent process planning framework of aviation parts // <https://doi.org/10.1080/00207543.2021.1951869>. Taylor & Francis, 2021. Т. 60, № 17. С. 5217–5234.
 109. Cao B. и др. A Multiobjective Intelligent Decision-Making Method for Multistage Placement of PMU in Power Grid Enterprises // IEEE Trans. Ind. Informatics. IEEE Computer Society, 2022.
 110. Xie J., Li S., Wang X. A digital smart product service system and a case study of the mining industry: MSPSS // Adv. Eng. Informatics. Elsevier, 2022. Т. 53. С. 101694.
 111. Dev S. и др. A predictive analytics approach for stroke prediction using machine learning and neural networks // Healthc. Anal. Elsevier, 2022. Т. 2. С. 100032.
 112. Zhang H. и др. Big data-assisted social media analytics for business model for business decision making system competitive analysis // Inf. Process. Manag. Pergamon, 2022. Т. 59, № 1. С. 102762.
 113. Yang J. и др. Social media data analytics for business decision making system to competitive analysis // Inf. Process. Manag. Pergamon, 2022. Т. 59, № 1. С. 102751.
 114. Singh S., Kumar A. Application of Quantum Computing Based Artificial Intelligence Algorithm for Business Development // J. Pharm. Negat. Results. 2022. Т. 13. С. 2060–2065.
 115. Ullah M.H. и др. Quantum computing for smart grid applications // IET Gener. Transm. Distrib. The Institution of Engineering and Technology, 2022. Т. 16, № 21. С. 4239–4257.
 116. RACHEL LEE. What is a Decision Support System in Artificial Intelligence (AI)? | Sisu Data [Электронный ресурс] // Sisu Data. 2021. URL: <https://sisudata.com/blog/what-is-a-decision-support-system-in-artificial-intelligence> (дата обращения: 18.06.2023).

117. Nauata N. и др. House-GAN: Relational Generative Adversarial Networks for Graph-Constrained House Layout Generation // *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. Springer Science and Business Media Deutschland GmbH, 2020. Т. 12346 LNCS. С. 162–177.
118. Lorusso A., Messina B., Santaniello D. The Use of Generative Adversarial Network as Graphical Support for Historical Urban Renovation // *Lect. Notes Data Eng. Commun. Technol.* Springer Science and Business Media Deutschland GmbH, 2023. Т. 146. С. 738–748.
119. Nobari A.H., Rashad M.F., Ahmed F. CreativeGAN: Editing Generative Adversarial Networks for Creative Design Synthesis // *Proc. ASME Des. Eng. Tech. Conf. American Society of Mechanical Engineers (ASME)*, 2021. Т. 3A-2021.
120. Jin L., Tan F., Jiang S. Generative Adversarial Network Technologies and Applications in Computer Vision // *Comput. Intell. Neurosci.* Hindawi Limited, 2020. Т. 2020.
121. Li T. и др. Analysis on Generative Adversarial Networks: Application in Graphics and Image Processing // *Proc. 2022 IEEE 4th Int. Conf. Civ. Aviat. Saf. Inf. Technol. ICCASIT 2022.* Institute of Electrical and Electronics Engineers Inc., 2022. С. 822–827.
122. Wang F. и др. An Adaptive Learning Image Denoising Algorithm Based on Eigenvalue Extraction and the GAN Model // *Comput. Intell. Neurosci.* Hindawi Limited, 2022. Т. 2022.
123. Duan Y., Zhang J. A Novel AI-Based Visual Stimuli Generation Approach for Environment Concept Design // *Comput. Intell. Neurosci.* Hindawi Limited, 2022. Т. 2022.
124. Mathesul S., Bhutkar G., Rambhad A. AttnGAN: Realistic Text-to-Image Synthesis with Attentional Generative Adversarial Networks // *Lect. Notes Comput. Sci. (including Subser. Lect. Notes Artif. Intell. Lect. Notes Bioinformatics)*. Springer Science and Business Media Deutschland GmbH, 2022. Т. 13198 LNCS. С. 397–403.
125. Zhang Y. и др. FHRGAN: Generative adversarial networks for synthetic fetal heart rate signal generation in low-resource settings // *Inf. Sci. (Ny)*. Elsevier, 2022. Т. 594. С. 136–150.
126. Zhong G. и др. Multi-Scale Attention Generative Adversarial Network for Medical Image Enhancement // *IEEE Trans. Emerg. Top. Comput. Intell.* Institute of Electrical and Electronics Engineers Inc., 2023.
127. Ali H., Shah Z. Combating COVID-19 Using Generative Adversarial Networks and Artificial Intelligence for Medical Images: Scoping Review // *JMIR Med Inf.* 2022;10(6)e37365 <https://medinform.jmir.org/2022/6/e37365>. *JMIR Medical Informatics*, 2022. Т. 10, № 6. С. e37365.
128. Staffini A. Stock Price Forecasting by a Deep Convolutional Generative Adversarial Network // *Front. Artif. Intell.* Frontiers Media S.A., 2022. Т. 5.

129. Han F., Ma X., Zhang J. Simulating Multi-Asset Classes Prices Using Wasserstein Generative Adversarial Network: A Study of Stocks, Futures and Cryptocurrency // *J. Risk Financ. Manag.* 2022, Vol. 15, Page 26. Multidisciplinary Digital Publishing Institute, 2022. Т. 15, № 1. С. 26.
130. Staffini A. Stock Price Forecasting by a Deep Convolutional Generative Adversarial Network // *Front. Artif. Intell. Frontiers Media S.A.*, 2022. Т. 5. С. 8.
131. Terzioğlu S., Çoğalmış K.N., Bulut A. Ad creative generation using reinforced generative adversarial network // *Electron. Commer. Res. Springer*, 2022. С. 1–17.
132. Polamuri S.R., Srinivas D.K., Krishna Mohan D.A. Multi-Model Generative Adversarial Network Hybrid Prediction Algorithm (MMGAN-HPA) for stock market prices prediction // *J. King Saud Univ. - Comput. Inf. Sci. Elsevier*, 2022. Т. 34, № 9. С. 7433–7444.
133. Uemura T. и др. A generative flow-based model for volumetric data augmentation in 3D deep learning for computed tomographic colonography // *Int. J. Comput. Assist. Radiol. Surg.* Springer Science and Business Media Deutschland GmbH, 2021. Т. 16, № 1. С. 81–89.
134. Revin I. и др. Automated machine learning approach for time series classification pipelines using evolutionary optimization // *Knowledge-Based Syst. Elsevier*, 2023. Т. 268. С. 110483.
135. Biswal S. и др. EMIXER: End-to-end Multimodal X-ray Generation via Self-supervision // *Proceedings of Machine Learning Research. PMLR*, 2022. Т. 182. С. 297–324.
136. McDuff D., Curran T., Kadambi A. Synthetic Data in Healthcare // *arxiv.org*. 2023.
137. Wang R., Chen Z., Li W. Gradient flow-based meta generative adversarial network for data augmentation in fault diagnosis // *Appl. Soft Comput. Elsevier*, 2023. Т. 142. С. 110313.
138. Zhang L. и др. End-to-end unsupervised fault detection using a flow-based model // *Reliab. Eng. Syst. Saf. Elsevier*, 2021. Т. 215. С. 107805.
139. Lin L. и др. Residential Electricity Load Scenario Prediction Based on Transferable Flow Generation Model // *J. Electr. Eng. Technol. Korean Institute of Electrical Engineers*, 2023. Т. 18, № 1. С. 99–109.
140. Burnap A., Hauser J.R., Timoshenko A. Design and Evaluation of Product Aesthetics: A Human-Machine Hybrid Approach // *SSRN Electron. J. Elsevier BV*, 2021.
141. Kim H.G. и др. Pricing path-dependent exotic options with flow-based generative networks // *Appl. Soft Comput. Elsevier*, 2022. Т. 124. С. 109049.
142. Cao M. и др. Recognition of Occluded Goods under Prior Inference Based on Generative Adversarial Network // *Sensors (Basel). NLM (Medline)*, 2023. Т. 23, № 6.
143. RICK MERRITT. What Is a Transformer Model? [Электронный ресурс] // *NVIDIA Blogs*. 2022. URL: <https://blogs.nvidia.com/blog/2022/03/25/what-is-a-transformer-model/> (дата обращения: 09.06.2023).
144. Uluslu A.Y., Schneider G. Scaling Native Language Identification with Transformer Adapters

- // arxiv.org. 2022.
145. Rivas P., Zhao L. Marketing with ChatGPT: Navigating the Ethical Terrain of GPT-Based Chatbot Technology // AI 2023. Multidisciplinary Digital Publishing Institute, 2023. Т. 4, № 2. С. 375–384.
 146. Ghourabi A., Alohaly M. Enhancing Spam Message Classification and Detection Using Transformer-Based Embedding and Ensemble Learning // Sensors. MDPI, 2023. Т. 23, № 8.
 147. Zhang A. и др. Shifting machine learning for healthcare from development to deployment and from models to data // Nat. Biomed. Eng. 2022 612. Nature Publishing Group, 2022. Т. 6, № 12. С. 1330–1345.
 148. Smith B., Van Steelandt S., Khojandi A. Evaluating the Impact of Health Care Data Completeness for Deep Generative Models // Methods Inf. Med. Georg Thieme Verlag KG, 2023. Т. 62, № 01/02. С. 031–039.
 149. Sezgin E., Sirrianni J., Linwood S.L. Operationalizing and Implementing Pretrained, Large Artificial Intelligence Linguistic Models in the US Health Care System: Outlook of Generative Pretrained Transformer 3 (GPT-3) as a Service Model // JMIR Med Inf. 2022;10(2)e32875 <https://medinform.jmir.org/2022/2/e32875>. JMIR Medical Informatics, 2022. Т. 10, № 2. С. e32875.
 150. Hirose T. и др. Diagnostic Accuracy of Differential-Diagnosis Lists Generated by Generative Pretrained Transformer 3 Chatbot for Clinical Vignettes with Common Chief Complaints: A Pilot Study // Int. J. Environ. Res. Public Health. MDPI, 2023. Т. 20, № 4. С. 3378.
 151. Beerbaum Dr. D.O. Generative Artificial Intelligence (GAI) with Chat GPT for Accounting – a Business Case // SSRN Electron. J. Elsevier BV, 2023.
 152. Hu N., Liang P., Yang X. Whetting All Your Appetites for Financial Tasks with One Meal from GPT: A Comparison of GPT, FinBERT, and Dictionaries in Evaluating Sentiment Analysis. 2023.
 153. Niszczoła P., Abbas S. GPT as a Financial Advisor // SSRN Electron. J. 2023.
 154. Suparna Biswas B.C., LastName S.S. AI in banking: Can banks meet the challenge? [Электронный ресурс] // McKinsey. 2020. URL: <https://www.mckinsey.com/industries/financial-services/our-insights/ai-bank-of-the-future-can-banks-meet-the-ai-challenge> (дата обращения: 09.06.2023).
 155. Liu D., Farajalla G.P., Boulenger A. Transformer-based Banking Products Recommender System // SIGIR 2021 - Proc. 44th Int. ACM SIGIR Conf. Res. Dev. Inf. Retr. Association for Computing Machinery, Inc, 2021. С. 2641–2642.
 156. Bender E.M. и др. On the dangers of stochastic parrots: Can language models be too big? //

- FAccT 2021 - Proc. 2021 ACM Conf. Fairness, Accountability, Transpar. Association for Computing Machinery, Inc, 2021. С. 610–623.
157. Korangi K., Mues C., Bravo C. A transformer-based model for default prediction in mid-cap corporate markets // *Eur. J. Oper. Res. Elsevier B.V.*, 2021. Т. 308, № 1. С. 306–320.
 158. Sharma U., Pandey P., Kumar S. A Transformer-Based Model for Evaluation of Information Relevance in Online Social-Media: A Case Study of Covid-19 Media Posts // *New Gener. Comput. Springer*, 2022. Т. 40, № 4. С. 1029–1052.
 159. Raj Dash, Andreas Kremer A., Petrov A. Designing next-generation credit-decisioning models [Электронный ресурс] // *McKinsey*. 2021. URL: <https://www.mckinsey.com/capabilities/risk-and-resilience/our-insights/designing-next-generation-credit-decisioning-models> (дата обращения: 09.06.2023).
 160. Hvatov A. DATA-DRIVEN APPROACH FOR THE FLOQUET PROPAGATOR INVERSE PROBLEM SOLUTION // *ICASSP, IEEE Int. Conf. Acoust. Speech Signal Process. - Proc. Institute of Electrical and Electronics Engineers Inc.*, 2022. Т. 2022-May. С. 3813–3817.
 161. Durai S.K.S., Shamili M.D. Smart farming using Machine Learning and Deep Learning techniques // *Decis. Anal. J. Elsevier*, 2022. Т. 3. С. 100041.
 162. Quazi S. Artificial intelligence and machine learning in precision and genomic medicine // *Med. Oncol.* 2022 398. *Springer*, 2022. Т. 39, № 8. С. 1–18.
 163. Bahrami M., Forouzanfar M. Sleep Apnea Detection from Single-Lead ECG: A Comprehensive Analysis of Machine Learning and Deep Learning Algorithms // *IEEE Trans. Instrum. Meas. Institute of Electrical and Electronics Engineers Inc.*, 2022. Т. 71.
 164. İnce M. Automatic and intelligent content visualization system based on deep learning and genetic algorithm // *Neural Comput. Appl. Springer Science and Business Media Deutschland GmbH*, 2022. Т. 34, № 3. С. 2473–2493.
 165. Ansari S. и др. Automatic Digital Modulation Recognition Based on Genetic-Algorithm-Optimized Machine Learning Models // *IEEE Access. Institute of Electrical and Electronics Engineers Inc.*, 2022. Т. 10. С. 50265–50277.
 166. Gaber A. и др. Automatic Classification of Fatty Liver Disease Based on Supervised Learning and Genetic Algorithm // *Appl. Sci.* 2022, Vol. 12, Page 521. *Multidisciplinary Digital Publishing Institute*, 2022. Т. 12, № 1. С. 521.
 167. Plakias S., Boutalis Y.S. A novel information processing method based on an ensemble of Auto-Encoders for unsupervised fault detection // *Comput. Ind. Elsevier*, 2022. Т. 142. С. 103743.
 168. Dong Z. и др. Data-driven power system reliability evaluation based on stacked denoising auto-encoders // *Energy Reports. Elsevier*, 2022. Т. 8. С. 920–927.

169. Maggipinto M., Beghi A., Susto G.A. A Deep Convolutional Autoencoder-Based Approach for Anomaly Detection with Industrial, Non-Images, 2-Dimensional Data: A Semiconductor Manufacturing Case Study // *IEEE Trans. Autom. Sci. Eng. Institute of Electrical and Electronics Engineers Inc.*, 2022. Т. 19, № 3. С. 1477–1490.
170. Rüttgers M. и др. A machine-learning-based method for automatizing lattice-Boltzmann simulations of respiratory flows // *Appl. Intell. Springer*, 2022. Т. 52, № 8. С. 9080–9100.
171. Yu G., Xiao L., Song W. Deep learning-based heterogeneous strategy for customizing responses of lattice structures // *Int. J. Mech. Sci. Pergamon*, 2022. Т. 229. С. 107531.
172. Jarin S. и др. Predicting the Crystal Structure and Lattice Parameters of the Perovskite Materials via Different Machine Learning Models Based on Basic Atom Properties // *Crystals. MDPI*, 2022. Т. 12, № 11. С. 1570.
173. Markovics D., Mayer M.J. Comparison of machine learning methods for photovoltaic power forecasting based on numerical weather prediction // *Renew. Sustain. Energy Rev. Pergamon*, 2022. Т. 161. С. 112364.
174. Liu X., Wu J., Chen S. A context-based meta-reinforcement learning approach to efficient hyperparameter optimization // *Neurocomputing. Elsevier*, 2022. Т. 478. С. 89–103.
175. Nikitin N.O. и др. Automated evolutionary approach for the design of composite machine learning pipelines // *Futur. Gener. Comput. Syst. North-Holland*, 2022. Т. 127. С. 109–125.
176. Altan A., Karasu S. Crude oil time series prediction model based on LSTM network with chaotic Henry gas solubility optimization // *Energy. Pergamon*, 2022. Т. 242. С. 122964.
177. Trevino-Martinez S., Sawhney R., Shylo O. Energy-carbon footprint optimization in sequence-dependent production scheduling // *Appl. Energy. Elsevier*, 2022. Т. 315. С. 118949.
178. Hertel L., Baldi P., Gillen D.L. Reproducible Hyperparameter Optimization // *Inf. UK Ltd. Taylor & Francis*, 2021. Т. 31, № 1. С. 84–99.
179. Lee J.H., Kim H.J. Reinforcement learning for robotic flow shop scheduling with processing time variations // *Inf. UK Ltd. Taylor & Francis*, 2021. Т. 60, № 7. С. 2346–2368.
180. Han Y. и др. Deep Reinforcement Learning for Robot Collision Avoidance With Self-State-Attention and Sensor Fusion // *IEEE Robot. Autom. Lett. Institute of Electrical and Electronics Engineers Inc.*, 2022. Т. 7, № 3. С. 6886–6893.
181. Thumm J., Althoff M. Provably Safe Deep Reinforcement Learning for Robotic Manipulation in Human Environments // *Proc. - IEEE Int. Conf. Robot. Autom. Institute of Electrical and Electronics Engineers Inc.*, 2022. С. 6344–6350.
182. Feng S. и др. Dense reinforcement learning for safety validation of autonomous vehicles // *Nat. 2023 6157953. Nature Publishing Group*, 2023. Т. 615, № 7953. С. 620–627.
183. Li G. и др. Decision making of autonomous vehicles in lane change scenarios: Deep

- reinforcement learning approaches with risk awareness // *Transp. Res. Part C Emerg. Technol.* Pergamon, 2022. T. 134. C. 103452.
184. Du Y. и др. Comfortable and energy-efficient speed control of autonomous vehicles on rough pavements using deep reinforcement learning // *Transp. Res. Part C Emerg. Technol.* Pergamon, 2022. T. 134. C. 103489.
 185. Alanazi A. Using machine learning for healthcare challenges and opportunities // *Informatics Med. Unlocked.* Elsevier, 2022. T. 30. C. 100924.
 186. Dai Y. и др. A closed-loop healthcare processing approach based on deep reinforcement learning // *Multimed. Tools Appl.* Springer, 2022. T. 81, № 3. C. 3107–3129.
 187. Guerrero J.I., Miró-Amarante G., Martín A. Decision support system in health care building design based on case-based reasoning and reinforcement learning // *Expert Syst. Appl.* Pergamon, 2022. T. 187. C. 116037.
 188. Zhang H. и др. H2GNN: Hierarchical-Hops Graph Neural Networks for Multi-Robot Exploration in Unknown Environments // *IEEE Robot. Autom. Lett.* Institute of Electrical and Electronics Engineers Inc., 2022. T. 7, № 2. C. 3435–3442.
 189. Gao P. и др. Improved Position Estimation Algorithm of Agricultural Mobile Robots Based on Multisensor Fusion and Autoencoder Neural Network // *Sensors 2022*, Vol. 22, Page 1522. Multidisciplinary Digital Publishing Institute, 2022. T. 22, № 4. C. 1522.
 190. Yu J., Vincent J.A., Schwager M. DiNNO: Distributed Neural Network Optimization for Multi-Robot Collaborative Learning // *IEEE Robot. Autom. Lett.* Institute of Electrical and Electronics Engineers Inc., 2022. T. 7, № 2. C. 1896–1903.
 191. Cheng D. и др. Financial time series forecasting with multi-modality graph neural network // *Pattern Recognit.* Pergamon, 2022. T. 121. C. 108218.
 192. Martínez F. и др. Strategies for time series forecasting with generalized regression neural networks // *Neurocomputing.* Elsevier, 2022. T. 491. C. 509–521.
 193. Wu J. и др. Price graphs: Utilizing the structural information of financial time series for stock prediction // *Inf. Sci. (Ny).* Elsevier, 2022. T. 588. C. 405–424.
 194. Patel R., Mandradia D., George K. A Novel Training Algorithm for Neuro-control of Discrete-time Nonlinear Systems // *2022 IEEE Int. Conf. Electron. Comput. Commun. Technol. CONECCT 2022.* Institute of Electrical and Electronics Engineers Inc., 2022.
 195. Xu M., Yang X. Reinforcement Learning for Robust Neuro-Control of Constrained Nonlinear Systems // *2022 4th Int. Conf. Data-Driven Optim. Complex Syst. DOCS 2022.* Institute of Electrical and Electronics Engineers Inc., 2022.
 196. Rego R.C.B., de Araújo F.M.U. Learning-based robust neuro-control: A method to compute control Lyapunov functions // *Int. J. Robust Nonlinear Control.* John Wiley & Sons, Ltd, 2022.

- T. 32, № 5. С. 2644–2661.
197. Maslyaev M., Hvatov A. Solver-Based Fitness Function for the Data-Driven Evolutionary Discovery of Partial Differential Equations // 2022 IEEE Congr. Evol. Comput. CEC 2022 - Conf. Proc. Institute of Electrical and Electronics Engineers Inc., 2022.
 198. Semenov A. и др. On Probabilistic Generalization of Backdoors in Boolean Satisfiability // Proc. AAAI Conf. Artif. Intell. Association for the Advancement of Artificial Intelligence, 2022. Т. 36, № 9. С. 10353–10361.
 199. Pavlenko A., Chivilikhin D., Semenov A. Asynchronous Evolutionary Algorithm for Finding Backdoors in Boolean Satisfiability // 2022 IEEE Congr. Evol. Comput. CEC 2022 - Conf. Proc. Institute of Electrical and Electronics Engineers Inc., 2022.
 200. Buzdalov M. The $(1+(\lambda,\lambda))$ Genetic Algorithm on the Vertex Cover Problem: Crossover Helps Leaving Plateaus // 2022 IEEE Congr. Evol. Comput. CEC 2022 - Conf. Proc. Institute of Electrical and Electronics Engineers Inc., 2022.
 201. Yan X. и др. Boundary consensus control strategies for fractional-order multi-agent systems with reaction-diffusion terms // Inf. Sci. (Ny). Elsevier, 2022. Т. 616. С. 461–473.
 202. Levshun D., Kotenko I. A survey on artificial intelligence techniques for security event correlation: models, challenges, and opportunities // Artif. Intell. Rev. Springer Nature, 2023. С. 1–44.
 203. Borisov I.I. и др. Reconfigurable Underactuated Adaptive Gripper Designed by Morphological Computation // Proc. - IEEE Int. Conf. Robot. Autom. Institute of Electrical and Electronics Engineers Inc., 2022. С. 1130–1136.
 204. Liu P. и др. Multi-agent based optimal equilibrium selection with resilience constraints for traffic flow // Neural Networks. Pergamon, 2022. Т. 155. С. 308–317.
 205. Agarwal A. и др. Stronger Generalization Guarantees for Robot Learning by Combining Generative Models and Real-World Data // Proc. - IEEE Int. Conf. Robot. Autom. Institute of Electrical and Electronics Engineers Inc., 2022. С. 4414–4421.
 206. Chen S. и др. Automatic weld type classification, tacked spot recognition and weld ROI determination for robotic welding based on modified YOLOv5 // Robot. Comput. Integr. Manuf. Pergamon, 2023. Т. 81. С. 102490.
 207. Holzinger A. и др. Information fusion as an integrative cross-cutting enabler to achieve robust, explainable, and trustworthy medical artificial intelligence // Inf. Fusion. Elsevier, 2022. Т. 79. С. 263–278.
 208. Azizi S. и др. Robust and Efficient Medical Imaging with Self-Supervision // arxiv.org. 2022.
 209. Bulanova N., Buzdalova A., Doerr C. Fast Re-Optimization of LeadingOnes with Frequent Changes // 2022 IEEE Congr. Evol. Comput. CEC 2022 - Conf. Proc. Institute of Electrical

- and Electronics Engineers Inc., 2022.
210. Ryumina E., Dresvyanskiy D., Karpov A. In search of a robust facial expressions recognition model: A large-scale visual cross-corpus study // *Neurocomputing*. Elsevier, 2022. T. 514. C. 435–450.
 211. Hu Y., Long G., Liu J. Real-time data transmission of cross-border e-commerce based on big data considering data complexity // *Proc. - 4th Int. Conf. Smart Syst. Inven. Technol. ICSSIT 2022*. Institute of Electrical and Electronics Engineers Inc., 2022. C. 1118–1122.
 212. Wang L., Wang Y. Supply chain financial service management system based on block chain IoT data sharing and edge computing // *Alexandria Eng. J.* Elsevier, 2022. T. 61, № 1. C. 147–158.
 213. Wu Q. Real-time Predictive Analysis of Loan Risk with Intelligent Monitoring and Machine Learning Technique // *2022 IEEE 4th Int. Conf. Power, Intell. Comput. Syst. ICPICS 2022*. Institute of Electrical and Electronics Engineers Inc., 2022. C. 852–856.
 214. Awotunde J.B. и др. Big Data Analytics of IoT-Based Cloud System Framework: Smart Healthcare Monitoring Systems // *Internet of Things*. Springer Science and Business Media Deutschland GmbH, 2022. C. 181–208.
 215. Guo K. и др. Federated Learning Empowered Real-Time Medical Data Processing Method for Smart Healthcare // *IEEE/ACM Trans. Comput. Biol. Bioinforma.* Institute of Electrical and Electronics Engineers Inc., 2022.
 216. Demirdöğen G., Işık Z., Arayıcı Y. BIM-based big data analytic system for healthcare facility management // *J. Build. Eng.* Elsevier, 2023. T. 64. C. 105713.
 217. Ma S. и др. Digital twin and big data-driven sustainable smart manufacturing based on information management systems for energy-intensive industries // *Appl. Energy*. Elsevier, 2022. T. 326. C. 119986.
 218. Hader M. и др. Applying integrated Blockchain and Big Data technologies to improve supply chain traceability and information sharing in the textile sector // *J. Ind. Inf. Integr.* Elsevier, 2022. T. 28. C. 100345.
 219. Mostafa N., Ramadan H.S.M., Elfarouk O. Renewable energy management in smart grids by using big data analytics and machine learning // *Mach. Learn. with Appl.* Elsevier, 2022. T. 9. C. 100363.
 220. EL Azzaoui A., Sharma P.K., Park J.H. Blockchain-based delegated Quantum Cloud architecture for medical big data security // *J. Netw. Comput. Appl.* Academic Press, 2022. T. 198. C. 103304.
 221. Sheffield N.C. и др. From biomedical cloud platforms to microservices: next steps in FAIR data and analysis // *Sci. Data 2022 91*. Nature Publishing Group, 2022. T. 9, № 1. C. 1–8.

222. Ahmad M. и др. Biomedical Applications Cloud-Based IoT // Biomed. Transl. Res. Springer, Singapore, 2022. С. 141–153.
223. Rehman A., Mehmood H. Influence of Big Data Analytics and Cloud Computing in the E-Commerce Industry // Int. J. Secur. Intell. Comput. (IJSIC). 2022. Т. 1, № 1. С. 41–53.
224. Chen J. Research on Smart Analysis System of E-commerce Trade Based on Computer Big Data and Cloud Computing // 2022 IEEE Int. Conf. Electr. Eng. Big Data Algorithms, EEBDA 2022. Institute of Electrical and Electronics Engineers Inc., 2022. С. 61–64.
225. Zeng Y. и др. E-Commerce Network Security Based on Big Data in Cloud Computing Environment // Mob. Inf. Syst. Hindawi Limited, 2022. Т. 2022.
226. Liu C. и др. Service-oriented industrial internet of things gateway for cloud manufacturing // Robot. Comput. Integr. Manuf. Pergamon, 2022. Т. 73. С. 102217.
227. Memos V.A., Psannis K.E., Lv Z. A Secure Network Model Against Bot Attacks in Edge-Enabled Industrial Internet of Things // IEEE Trans. Ind. Informatics. IEEE Computer Society, 2022. Т. 18, № 11. С. 7998–8006.
228. Wei D. и др. Personalized Privacy-Aware Task Offloading for Edge-Cloud-Assisted Industrial Internet of Things in Automated Manufacturing // IEEE Trans. Ind. Informatics. IEEE Computer Society, 2022. Т. 18, № 11. С. 7935–7945.
229. Dahiya A. и др. A comprehensive analysis of blockchain and its applications in intelligent systems based on IoT, cloud and social media // Int. J. Intell. Syst. John Wiley & Sons, Ltd, 2022. Т. 37, № 12. С. 11037–11077.
230. Muheidat F. и др. Emerging Concepts Using Blockchain and Big Data // Procedia Comput. Sci. Elsevier, 2022. Т. 198. С. 15–22.
231. Gutiérrez Torre A. Distributed cloud-edge analytics and machine learning for transportation emissions estimation // TDX (Tesis Dr. en Xarxa). Universitat Politècnica de Catalunya, 2022.
232. Thirumahal R., Sudha Sadasivam G., Shruti P. Semantic Integration of Heterogeneous Data Sources Using Ontology-Based Domain Knowledge Modeling for Early Detection of COVID-19 // SN Comput. Sci. Springer, 2022. Т. 3, № 6.
233. Martínez-García M., Hernández-Lemus E. Data Integration Challenges for Machine Learning in Precision Medicine // Front. Med. Frontiers Media S.A., 2022. Т. 8.
234. Saggi M.K., Jain S. A Survey Towards Decision Support System on Smart Irrigation Scheduling Using Machine Learning approaches // Arch. Comput. Methods Eng. Springer Science and Business Media B.V., 2022. С. 1–24.
235. Wang S., Qin A. Design and Implementation of a Multidimensional Visualization Reconstruction System for Old Urban Spaces Based on Neural Networks // Comput. Intell. Neurosci. Hindawi Limited, 2022. Т. 2022.

236. Sun C., Wang K.C. DLA-VPS: Deep-Learning-Assisted Visual Parameter Space Analysis of Cosmological Simulations // IEEE Comput. Graph. Appl. IEEE Computer Society, 2022. Т. 42, № 3. С. 41–52.
237. Junaid M. и др. Performance Evaluation of Data-driven Intelligent Algorithms for Big data Ecosystem // Wirel. Pers. Commun. Springer, 2022. Т. 126, № 3. С. 2403–2423.
238. Venkateswarlu Y. и др. An Efficient Outlier Detection with Deep Learning-Based Financial Crisis Prediction Model in Big Data Environment // Comput. Intell. Neurosci. Hindawi Limited, 2022. Т. 2022.
239. Lee I., Mangalaraj G. Big Data Analytics in Supply Chain Management: A Systematic Literature Review and Research Directions // Big Data Cogn. Comput. 2022, Vol. 6, Page 17. Multidisciplinary Digital Publishing Institute, 2022. Т. 6, № 1. С. 17.
240. Zeng Z. и др. Statistical and machine learning methods for spatially resolved transcriptomics data analysis // Genome Biol. BioMed Central Ltd, 2022. Т. 23, № 1. С. 1–23.
241. Chiranjil Lal Chowdhary P., G. Thippa Reddy P., B. D. Parameshachari P. Computer vision and recognition systems : research innovations and trends // Apple Acad. Press Inc. 2022.
242. Rizzoli A. 27+ Most Popular Computer Vision Applications in 2023 [Электронный ресурс] // V7 Labs. 2023. URL: <https://www.v7labs.com/blog/computer-vision-applications> (дата обращения: 30.04.2023).
243. Abhijith N. Computer Vision Market Size, Trends | Industry Forecast - 2030 [Электронный ресурс] // Allied Market Research. URL: <https://www.alliedmarketresearch.com/computer-vision-market-A12701> (дата обращения: 30.04.2023).
244. Artificial Intelligence (AI) in Computer Vision Market | 2021-28| Size, Share Report 2028 [Электронный ресурс] // Polaris Market Research and Consulting. URL: <https://www.polarismarketresearch.com/industry-analysis/artificial-intelligence-ai-in-computer-vision-market> (дата обращения: 30.04.2023).
245. Яков Шпунт. Ретейлеры в лидерах использования ИИ [Электронный ресурс] // ComNews. 2023. URL: <https://www.comnews.ru/content/224549/2023-02-22/2023-w08/reteylery-liderakh-ispolzovaniya-ii> (дата обращения: 30.04.2023).
246. Николай Носов. Тенденции развития компьютерного зрения в России [Электронный ресурс] // IKSMEDIA. 2022. URL: <https://www.iksmedia.ru/articles/5894791-Tendencii-razvitiya-kompyuternogo.html> (дата обращения: 30.04.2023).
247. Сбер представил набор инструментов для встраивания систем компьютерного зрения в мобильные приложения [Электронный ресурс] // Сбер. 2023. URL: <https://sberlabs.com/news/sber-predstavil-nabor-instrumentov-dlya-vstraivaniya-sistem-kompyuternogo-zreniya-v-mobilnye-prilozh/> (дата обращения: 01.05.2023).

248. SberShuttle [Электронный ресурс] // tadviser. 2022. URL: <https://www.tadviser.ru/index.php/Продукт:SberShuttle> (дата обращения: 01.05.2023).
249. Natural Language Processing Market Forecast 2030 [Электронный ресурс] // Fortune Business Insights. 2022. URL: <https://www.fortunebusinessinsights.com/industry-reports/natural-language-processing-nlp-market-101933> (дата обращения: 24.05.2023).
250. Natural Language Processing Market Growth Report, 2030 [Электронный ресурс] // Grand View Research. 2022. URL: <https://www.grandviewresearch.com/industry-analysis/natural-language-processing-market-report> (дата обращения: 24.05.2023).
251. Исследование. Рынок разговорного ИИ в России 2020-2025 | Умные чат-боты и голосовые ассистенты для сложных бизнес-задач [Электронный ресурс] // Just AI. 2021. URL: <https://just-ai.com/blog/issledovanie-rynok-razgovornogo-ii-v-rossii-2020-2025> (дата обращения: 25.05.2023).
252. Voice And Speech Recognition Market Size Report, 2030 [Электронный ресурс] // Grand View Research. 2022. URL: <https://www.grandviewresearch.com/industry-analysis/voice-recognition-market> (дата обращения: 25.05.2023).
253. Кирилл Петров. Разговорный ИИ в 2022 году: взросление голосовых ассистентов, тренд мобильных секретарей и рост спроса на умные устройства [Электронный ресурс] // ICT.Moscow. 2023. URL: <https://ict.moscow/news/2022-conversational-ai/> (дата обращения: 26.05.2023).
254. Straits Research. Recommendation Engines Market Size is projected to reach [Электронный ресурс] // GlobeNewswire. 2022. URL: <https://www.globenewswire.com/en/news-release/2022/07/07/2476198/0/en/Recommendation-Engines-Market-Size-is-projected-to-reach-USD-54-Billion-by-2030-grow-at-a-CAGR-of-37-Straits-Research.html> (дата обращения: 27.05.2023).
255. Business Intelligence (рынок России) [Электронный ресурс] // tadviser. 2022. URL: [https://www.tadviser.ru/index.php/Статья:Business_Intelligence_\(рынок_России\)#.D0.9F.D0.BE.D0.B4.D1.8A.D0.B5.D0.BC_2021_.D0.B3.D0.BE.D0.B4.D0.B0](https://www.tadviser.ru/index.php/Статья:Business_Intelligence_(рынок_России)#.D0.9F.D0.BE.D0.B4.D1.8A.D0.B5.D0.BC_2021_.D0.B3.D0.BE.D0.B4.D0.B0) (дата обращения: 27.05.2023).
256. Tanmay S M.T. Generative AI Market Size, Share | Industry Report - 2031 [Электронный ресурс] // Allied Market Research. 2023. URL: <https://www.alliedmarketresearch.com/generative-ai-market-A47396> (дата обращения: 30.05.2023).
257. МФТИ DeepPavlov - библиотека диалоговых систем - CNews [Электронный ресурс] // cnews. 2023. URL: https://www.cnews.ru/book/МФТИ_DeepPavlov_-_библиотека_диалоговых_систем (дата обращения: 01.06.2023).

258. Сбер. Автоматическое машинное обучение: как оно работает, при чём здесь ламы и где учиться его использовать — Машинное обучение [Электронный ресурс] // vc.ru. 2021. URL: <https://vc.ru/ml/296358-avtomaticheskoe-mashinnoe-obuchenie-kak-ono-rabotaet-pri-chem-zdes-lamy-i-gde-uchitsya-ego-ispolzovat> (дата обращения: 02.06.2023).
259. Сбербанк Auto ML [Электронный ресурс] // tadviser. 2019. URL: https://www.tadviser.ru/index.php/Продукт:Сбербанк_Auto_ML (дата обращения: 02.06.2023).
260. Big Data Analytics Market Size, Share | Growth Statistics [2030] [Электронный ресурс] // Fortune Business Insights. 2022. URL: <https://www.fortunebusinessinsights.com/big-data-analytics-market-106179> (дата обращения: 02.06.2023).
261. Большие данные прошли переоценку – Коммерсантъ [Электронный ресурс] // Kommersant. 2023. URL: <https://www.kommersant.ru/doc/5939856> (дата обращения: 02.06.2023).
262. Интернет вещей, IoT, M2M (рынок России) [Электронный ресурс] // TAdviser. 2022. URL: [https://www.tadviser.ru/index.php/Статья:Интернет_вещей,_IoT,_M2M_\(рынок_России\)](https://www.tadviser.ru/index.php/Статья:Интернет_вещей,_IoT,_M2M_(рынок_России)) (дата обращения: 29.06.2023).
263. Industrial IoT (IIoT) Market Size, Share, Trend & Growth Driver 2030 [Электронный ресурс] // MarketsandMarkets. 2023. URL: <https://www.marketsandmarkets.com/Market-Reports/industrial-internet-of-things-market-129733727.html> (дата обращения: 29.06.2023).
264. Объём рынка интернета вещей в России вырос на 18,3 % [Электронный ресурс] // Российская газета. 2021. URL: <https://rg.ru/2021/07/28/obem-rynka-interneta-veshchej-v-rossii-vyros-na-183.html> (дата обращения: 29.06.2023).
265. Перспективы развития беспроводных технологий на ближайшие пять лет [Электронный ресурс]. 2022. URL: <https://mcs.mail.ru/blog/10-perspektivnych-besprovodnych> (дата обращения: 29.06.2023).
266. Industry Report Analysis, Trends & Growth Drivers - 2030 [Электронный ресурс] // Wireless Connectivity Market Size. 2022. URL: <https://www.marketsandmarkets.com/Market-Reports/wireless-connectivity-market-192605963.html> (дата обращения: 29.06.2023).
267. К 2025 году сети 5G могут охватить до 80 % населения России [Электронный ресурс] // VKCloud. 2022. URL: <https://mcs.mail.ru/blog/prognoz-razvitiya-seti-5g-i-cifrovoj-ehkonomiki-ot-gsma> (дата обращения: 29.06.2023).
268. Global Electronic Chemicals and Materials Market Report [Электронный ресурс]. 2022. URL: <https://www.globenewswire.com/news-release/2023/06/28/2695933/0/en/Global-Electronic-Chemicals-and-Materials-Market-Report-2023-Demand-for-Semiconductors-to>

- Reach-Up-to-1-Trillion-by-2030-Creating-Lucrative-Opportunities.html (дата обращения: 29.06.2023).
269. Инженерное программное обеспечение (рынок России) [Электронный ресурс]. 2022. URL: [https://www.tadviser.ru/index.php/Статья:Инженерное_программное_обеспечение_\(рынок_России\)#.D0.A0.D0.BE.D1.81.D1.81.D0.B8.D0.B9.D1.81.D0.BA.D0.B8.D0.B5_.D1.80.D0.B0.D0.B7.D1.80.D0.B0.D0.B1.D0.BE.D1.82.D1.87.D0.B8.D0.BA.D0.B8_.D0.A1.D0.90.D0.9F.D0.A0_.D0.B](https://www.tadviser.ru/index.php/Статья:Инженерное_программное_обеспечение_(рынок_России)#.D0.A0.D0.BE.D1.81.D1.81.D0.B8.D0.B9.D1.81.D0.BA.D0.B8.D0.B5_.D1.80.D0.B0.D0.B7.D1.80.D0.B0.D0.B1.D0.BE.D1.82.D1.87.D0.B8.D0.BA.D0.B8_.D0.A1.D0.90.D0.9F.D0.A0_.D0.B) (дата обращения: 29.06.2023).
270. Computer Aided Design Market Size Global Report, 2022 - 2030 [Электронный ресурс]. 2022. URL: <https://www.polarismarketresearch.com/industry-analysis/computer-aided-design-cad-market> (дата обращения: 29.06.2023).
271. Черкесов Е. В России не хватает более 2000 отечественных САПР. Программисты получают колоссальный объём работ на ближайшие семь лет [Электронный ресурс] // CNews. 2022. URL: https://www.cnews.ru/news/top/2022-12-20_v_rossii_ne_hvataet_bole (дата обращения: 29.06.2023).
272. Автоматизированные помощники. Как развивается российский рынок робототехники? [Электронный ресурс]. 2023. URL: https://spbvedomosti.ru/news/country_and_world/avtomatizirovannye-pomoshchniki-kak-razvivaetsya-rossiyskiy-rynok-robototekhniki/ (дата обращения: 29.06.2023).
273. Industrial Robotics Market Value to Hit US\$ 142.8 Billion [Электронный ресурс]. 2023. URL: <https://www.globenewswire.com/en/news-release/2023/03/09/2624175/0/en/Industrial-Robotics-Market-Value-to-Hit-US-142-8-Billion-in-2032-At-CAGR-11-4.html> (дата обращения: 29.06.2023).
274. Gates B. Responding to Covid-19 — A Once-in-a-Century Pandemic? // N. Engl. J. Med. Massachusetts Medical Society, 2020. Т. 382, № 18. С. 1677–1679.