

РУССКИЙ ЯЗЫК В РАЗРАБОТКЕ ИСКУССТВЕННОГО ИНТЕЛЛЕКТА

Куанышев Б.Н., Мирзеев Х.А.

*Белорусский государственный университет информатики и радиоэлектроники
г. Минск, Республика Беларусь*

Арцыменя Д.Ф. – ст. преподаватель

В статье рассматривается использование русского языка в разработке искусственного интеллекта (ИИ), исследуется история и основные этапы в развитие ИИ на русском языке, описываются принципы работы ИИ с русским языком, приводятся примеры практического применения ИИ на русском языке.

Русский язык, благодаря своей богатой морфологии, сложной грамматике и широкой семантической вариативности, представляет собой уникальный объект для разработки технологий искусственного интеллекта (ИИ). В последние десятилетия стремительное развитие ИИ привело к созданию мощных языковых моделей и систем обработки текста, способных эффективно работать с русским языком. Современные исследования сосредоточены на решении задач, связанных с синтаксическим анализом, обработкой контекста и распознаванием многозначности, что делает

русскоязычные системы ИИ актуальными в различных областях: от автоматического перевода до создания виртуальных ассистентов. Настоящая статья рассматривает ключевые этапы развития ИИ для русского языка, основные принципы его работы и практические направления применения технологий.

Развитие искусственного интеллекта (ИИ) в контексте русского языка началось с середины XX века. Первыми шагами стало создание алгоритмов для перевода текстов с одного языка на другой. Советские ученые активно работали над системами машинного перевода, которые стали основой для дальнейших разработок.

С появлением компьютеров и развитием вычислительной техники в 70-80-х годах стало возможным создание более сложных систем обработки текстов. В это время разрабатывались программы, способные анализировать синтаксис и морфологию русского языка.

На рубеже XXI века начался переход к использованию статистических методов, которые позволили значительно улучшить точность обработки текста. В последние годы широкое распространение получили нейронные сети, обученные на огромных объемах данных, таких как языковые модели GPT и BERT, что сделало обработку русского языка более гибкой и точной [1, 2].

Русский язык представляет уникальные вызовы для разработчиков ИИ из-за его грамматической сложности. Например, использование падежей, богатый словарный запас, а также высокая степень контекстуальной зависимости требуют особого подхода.

Современные модели обработки естественного языка (NLP) работают с большими лингвистическими корпусами, содержащими миллионы текстов на русском языке. Нейронные сети анализируют данные и «учатся» различать контексты, строить правильные синтаксические конструкции и понимать значение слов в зависимости от их окружения.

Особенности обработки русского языка включают:

- морфологический анализ: определение основы и суффиксов слов;
- синтаксический анализ: выявление связей между словами;
- работа с неоднозначностями, такими как многозначные слова.

Для улучшения качества работы моделей ИИ применяются также дополнительные методы, такие как распознавание устной речи и обработка сленговых выражений [3, 4].

Современные технологии ИИ широко используются в различных сферах:

- машинный перевод: языковые модели обеспечивают быстрый и точный перевод текстов с русского на другие языки и наоборот;
- виртуальные ассистенты: голосовые помощники, такие как «Алиса» или «Сбер», активно используют русский язык для взаимодействия с пользователями;
- текстовый анализ: ИИ помогает анализировать большие объемы текстов, выявлять ключевые идеи и составлять отчеты;
- образование: с помощью ИИ разрабатываются приложения для изучения русского языка, включая интерактивные задания и персонализированные программы обучения;
- бизнес: использование чат-ботов и систем автоматической обработки данных позволяет компаниям работать эффективнее [5].

Развитие ИИ в контексте русского языка прошло долгий путь от первых попыток машинного перевода до создания сложных языковых моделей. Благодаря современным технологиям русский язык стал интегрированной частью глобального цифрового пространства. Несмотря на вызовы, связанные с его сложностью, принципы работы нейронных сетей и применение больших данных позволяют достигать высокой точности обработки текста. Практическое использование ИИ на русском языке открывает огромные перспективы, помогая развивать науку, бизнес и образование. В будущем такие технологии продолжат совершенствоваться, играя важную роль в русскоязычном обществе.

Список использованных источников:

1. Богуславский, И.М. Синтаксический анализ текстов на русском языке: проблемы и перспективы / И.М. Богуславский // Вопросы компьютерной лингвистики. – 2019. – № 3. – С. 12–25.
2. Иванов, А.А. Нейросетевые модели для обработки естественного языка: применение к русскому языку / А.А. Иванов // Программирование и компьютерные технологии. – 2020. – Т. 46, № 6. – С. 45–59.
3. Кукушкин, М.Н. Эволюция технологий машинного перевода: от статистических методов к нейронным сетям / М.Н. Кукушкин // Информационные технологии и вычислительные системы. – 2021. – № 5. – С. 33–47.
4. Федоров, С.Г. Разработка и использование лингвистических корпусов для русского языка в задачах NLP / С.Г. Фёдоров // Труды Российской академии наук. Информатика. – 2022. – Т. 55, № 4. – С. 67–81.
5. Сидоренко, Л.В. Применение искусственного интеллекта в анализе текстов и построении чат-ботов на русском языке / Л.В. Сидоренко // Наука и технологии в информационном обществе. – 2023. – Т. 10, № 2. – С. 95–109.