

УДК 004.8

ПРИМЕНЕНИЕ НЕЙРОСЕТЕВЫХ МОДЕЛЕЙ ДЛЯ ОБРАБОТКИ ЕСТЕСТВЕННОГО ЯЗЫКА

Монтик Николай Сергеевич, старший преподаватель
Манцевич Руслан Сергеевич, ассистент
Брестский государственный технический университет

APPLICATION OF NEURAL NETWORK MODELS TO NATURAL LANGUAGE PROCESSING

Montik Nikolay, Senior lecturer, nikolay.montik@gmail.com
Mantsevich Ruslan, assistant, bmw.dimon1994@gmail.com
Brest State Technical University

Аннотация. Рассматриваются нейросетевые подходы к NLP. Анализируются RNN, LSTM, CNN, BERT, GPT. Описываются машинный перевод, генерация текста, анализ тональности. Выделяются преимущества и ограничения нейросетевых моделей.

Ключевые слова: нейронные сети, обработка естественного языка, NLP, трансформеры, BERT, GPT, машинное обучение, компьютерная лингвистика.

Abstract. Neural network approaches to NLP are considered. RNN, LSTM, CNN, BERT, GPT are analyzed. Machine translation, text generation, sentiment analysis are described. Advantages and limitations of neural network models are highlighted.

Keywords: neural networks, natural language processing, NLP, transformers, BERT, GPT, machine learning, computational linguistics.

Введение. В последние десятилетия искусственный интеллект развивается очень быстро. Благодаря этому в области компьютерной лингвистики возникли принципиально иные методы. Язык человека — это не просто средство общения. Он представляет собой сложную систему, которая включает в себя лексику, грамматику, синтаксис и смысловые связи. Старые алгоритмические подходы строились на жестких правилах и лингвистических шаблонах. Они оказались недостаточно гибкими при работе с большими объёмами текстовых данных.

С появлением нейросетей и технологий глубокого обучения ситуация кардинально изменилась. Исследователи обучают модели на огромных текстовых корпусах. В результате нейросети научились выявлять скрытые зависимости между словами и даже генерировать связные тексты. Такие архитектуры, как RNN, LSTM, а затем трансформеры (BERT, GPT), стали фундаментом для современных NLP-систем.

В этом докладе рассматриваются главные направления в области обработки естественного языка с использованием нейросетей. Также анализируются их эффективность, проблемы интерпретируемости и дальнейшие перспективы.

Понятие и основные задачи обработки естественного языка

Обработка естественного языка (Natural Language Processing — NLP) представляет собой междисциплинарную область научного знания, интегрирующую лингвистику, информатику и искусственный интеллект. Целью NLP является разработка алгоритмических и модельных средств, обеспечивающих компьютерную интерпретацию, анализ и генерацию человеческой речи в текстовой либо звуковой форме.

В рамках NLP традиционно выделяется ряд ключевых задач. К их числу относится токенизация, под которой понимается сегментация текстового потока на дискретные элементы (токены) — слова, знаки препинания и морфемы. Морфологический анализ предполагает определение грамматических характеристик лексических единиц, а именно части речи, падежа, числа и других категорий. Синтаксический анализ заключается в построении грамматической структуры предложения и

выявлении синтаксических связей между его элементами. Семантический анализ направлен на извлечение смыслового содержания текста, установление значений слов и выражений с учётом контекстуальных факторов. Референциальная интерпретация обеспечивает установление анафорических связей между элементами текста (например, идентификацию референта местоимения «он»). Извлечение информации предполагает автоматическое выявление фактов, событий, именованных сущностей (людей, организаций, географических наименований) из неструктурированных текстовых массивов. Наконец, классификация текста заключается в отнесении текстового документа к определённой тематической или оценочной категории (например, спам или легитимное сообщение, позитивный или негативный отзыв).

На протяжении длительного периода перечисленные задачи решались на основе формальных лингвистических правил и статистических моделей. Однако указанные подходы характеризовались высокой трудоёмкостью, низкой адаптируемостью к новым языкам и недостаточной устойчивостью к лексической вариативности. Применение нейросетевых моделей позволило преодолеть отмеченные ограничения. Вместо ручной формализации правил модель обучается на больших текстовых корпусах, самостоятельно выявляя статистические закономерности. Данное обстоятельство обеспечило повышение точности и масштабируемости NLP-систем, что сделало возможным их практическое внедрение в поисковых системах, виртуальных ассистентах и иных прикладных областях.

Нейросетевые архитектуры, применяемые в NLP

Рекуррентные нейронные сети (Recurrent Neural Networks — RNN) явились первым классом моделей, продемонстрировавшим способность к обработке последовательных данных. Благодаря наличию циклических связей в архитектуре RNN обеспечивается запоминание информации о предшествующих элементах входной последовательности, что необходимо для учёта контекста при анализе текста. Однако классические RNN обладают существенным недостатком — проблемой исчезающего градиента, вследствие которой информация о начальных элементах длинных последовательностей утрачивается.

Для преодоления указанного ограничения были разработаны усовершенствованные архитектуры — LSTM (Long Short-Term Memory) и GRU (Gated Recurrent Unit). Данные модели характеризуются более устойчивым сохранением информации на протяжении протяжённых последовательностей. LSTM и GRU продемонстрировали высокую эффективность при решении задач машинного перевода, генерации текста и анализа тональности.

Свёрточные нейронные сети (Convolutional Neural Networks — CNN), получившие широкое распространение преимущественно в области компьютерного зрения, находят также применение в задачах NLP. CNN используются для извлечения локальных признаков из текстовых данных, что особенно актуально в контексте классификации текстов. Данные архитектуры эффективно выявляют локальные зависимости между словами (например, биграммы и триграммы), что полезно при анализе синтаксической структуры предложений.

Значительным прорывом в области NLP стало появление архитектуры Transformer, предложенной в 2017 году. Основу данной модели составляет механизм внимания, позволяющий учитывать контекст всего предложения без необходимости последовательной обработки, характерной для RNN. Transformer заменяет рекуррентные элементы на параллельные вычисления, что существенно повышает эффективность обучения и масштабируемость модели.

На основе архитектуры Transformer были разработаны мощные предобученные языковые модели. BERT обучается на задаче восстановления случайно скрытых слов в тексте, учитывая контекст как слева, так и справа от целевого токена. Данная модель широко применяется в задачах классификации, аннотирования текста и извлечения информации. GPT представляет собой авторегрессионную модель, предсказывающую следующее слово на основе предшествующего контекста. Она характеризуется способностью к генерации связных текстов и ведению диалога. Помимо указанных, существуют иные известные модели — RoBERTa, XLNet, T5, а также диалоговые системы семейства ChatGPT. Универсальность перечисленных моделей заключается в возможности их дообучения (fine-tuning) для решения конкретных прикладных задач без необходимости создания модели «с нуля», что обеспечивает экономическую эффективность нейросетевых подходов.

Преимущества и ограничения нейросетевых моделей

Применение нейросетевых подходов в NLP предоставляет ряд существенных преимуществ. К их числу относится универсальность — возможность адаптации единой архитектуры к различным прикладным задачам (от перевода до анализа тональности). Нейросетевые модели характеризуются масштабируемостью и высокой точностью, особенно при наличии больших объёмов обучающих данных. Кроме того, использование предобученных моделей позволяет существенно сократить вычислительные и временные ресурсы, необходимые для разработки специализированных решений, поскольку требуется лишь минимальная донастройка (fine-tuning).

Наряду с отмеченными преимуществами нейросетевые подходы обладают и серьёзными ограничениями. Одной из ключевых проблем остаётся низкая интерпретируемость: модели функционируют в режиме «чёрного ящика», что затрудняет объяснение причин принятия того или иного решения. Обучение крупных моделей требует значительных вычислительных мощностей и энергетических затрат, что ограничивает доступность данных технологий для индивидуальных пользователей и небольших организаций. Нейросети чувствительны к качеству обучающих данных: при наличии шумов, предвзятости или неравномерного распределения примеров возможны ошибки, выявление которых на этапе предварительного анализа затруднительно. Наконец, существуют этические риски, связанные с воспроизведением моделями предвзятых суждений, нарушением конфиденциальности данных и потенциальным использованием технологий во вред при отсутствии соответствующих механизмов контроля и регулирования.

Заключение. Развитие нейросетевых технологий привело к кардинальной трансформации методологических подходов к обработке естественного языка. Задачи, ранее требовавшие глубоких лингвистических знаний и значительных трудозатрат, в настоящее время решаются автоматически с использованием моделей глубокого обучения. Современные нейросетевые архитектуры демонстрируют высокую эффективность в машинном переводе, генерации текста, распознавании речи, классификации и извлечении информации. Универсальность, способность к обучению на больших объёмах данных и гибкость адаптации указанных моделей открывают широкие перспективы для их применения в научных исследованиях, бизнес-практике и повседневной деятельности.

Вместе с тем сохраняются серьёзные вызовы, включая проблему интерпретируемости, высокую ресурсоёмкость, этические риски и уязвимость к искажённым или предвзятым данным. Дальнейшее развитие данной области связывается с созданием более прозрачных, устойчивых и этически регулируемых моделей, способных к эффективному взаимодействию с человеком и учёту контекста в широком смысле.

Список использованных источников

1. Умаров А.А. Обработка естественного языка с использованием нейросетей / А.А. Умаров // *Innovations in Science and Technologies*. – 2025. – № 2. – С. 359–363.
2. Rogers A. A Primer in BERTology: What We Know about How BERT Works / A. Rogers, O. Kovaleva, A. Rumshisky // *Transactions of the Association for Computational Linguistics*. – 2020. – Vol. 8. – P. 842–866.
3. Jurafsky D. *Speech and Language Processing* / D. Jurafsky, J.H. Martin. – 3rd ed. – Boston: Pearson, 2023. – 1024 p.
4. Сапогов Е.М. Нейросетевые технологии в задачах NLP / Е.М. Сапогов. – М.: Наука, 2023. – 280 с.
5. Vaswani A. Attention is All You Need / A. Vaswani, N. Shazeer, N. Parmar et al. // *Advances in Neural Information Processing Systems*. – 2017. – Vol. 30. – P. 5998–6008.