

Селезнева Ольга Николаевна

канд. филол. наук, доцент

ФГБОУ ВО «Тверской государственной медицинский университет»

г. Тверь, Тверская область

DOI 10.31483/r-110426

ИСПОЛЬЗОВАНИЕ НЕЙРОСЕТЕЙ В ЛИНГВИСТИКЕ

***Аннотация:** на современном этапе технологического прогресса нейросети способны решать задачи высокого уровня сложности в различных сферах науки, бизнеса и образования. Особый интерес для лингвистики представляет технология обработки естественного языка (Natural Language Processing, NLP), которая дает возможность компьютерам интерпретировать человеческий язык.*

Нейросети могут использоваться в лингвистике: для написания научно-исследовательских работ, формулировки на родном и иностранном языках обучающих заданий, создания литературных и стихотворных произведений, выполнения быстрого и качественного перевода на иностранные языки, обучения и практики иностранного языка.

***Ключевые слова:** нейросеть, лингвистика, обработка естественного языка, NLP, иностранный язык, текст, машинный перевод, обучение иностранному языку, практика иностранного языка.*

Нейросети занимают все большее место в жизни современного человека, поскольку они зачастую способны решить задачи высокого уровня сложности в различных сферах науки, бизнеса и образования. Нейросеть представляет собой искусственный интеллект, при котором компьютерная программа имитирует, копирует работу мозга человека. Нейроны, которые находятся в мозге, передают друг другу сигналы, в то время как в нейросети информацией обмениваются вычислительные элементы.

Если обратиться к истории, то первая обучаемая одноуровневая нейросеть была продемонстрирована в 1957 году ученым, психологом по имени Фрэнк

Розенблатт [2]. Через 30 лет появились более мощные вычислительные компьютеры, что позволило исследователям разработать двух-, трехуровневые нейросети. Мощное развитие индустрии современных компьютерных игр произвели настоящую революцию в нейросетях, поскольку возникла потребность в сложных вычислениях для обработки огромного количества операций. Сложилось понимание, что архитектура графического процессора (GPU), который объединяет тысячи довольно простых вычислительных ядер на одном чипе, схожа с архитектурой нейросети [2]. В настоящее время развиваются самообучаемые нейросети, не требующие отдельных настроек, способные независимо анализировать входящую информацию. Искусственные нейроны, которые представляют собой программные модули или узлы, взаимосвязанные и взаимодействующие между собой, имитируют работу человеческих нейронов.

Базовая нейронная сеть включает в себя три уровня искусственных нейронов:

- 1) входной уровень, который обрабатывает внешнюю информацию, анализирует и «разбивает» ее, чтобы в дальнейшем передать на следующий уровень;
- 2) скрытый уровень, который классифицирует выходную информацию предыдущего уровня, анализирует и обрабатывает ее с целью передачи на следующий уровень;
- 3) выходной уровень, при котором происходит конечная обработка всех полученных данных и выдается окончательный результат [1; 2].

Примеры использования, применения нейросетей в науке, бизнесе, образовании многочисленны, но в рамках данной статьи рассмотрим использование нейросетей в различных областях лингвистики, что обусловлено их способностью анализировать и обрабатывать языковые данные. В этой связи необходимо обратиться к такому термину, как «обработка естественного языка (Natural Language Processing, NLP)». Согласно определению, «обработка естественного языка (NLP) – это технология машинного обучения, которая дает компьютерам возможность интерпретировать человеческий язык» [3]. Следует отметить, что во многих случаях NLP помогает эффективно анализировать как текстовые, так

и речевые данные. Применение нейросетей в различных областях лингвистики находит отражение в следующем:

1) обработка и анализ с помощью нейросетей различных текстовых заданий, а также генерирование текстов;

Работа нейросети для генерирования текстов выстроена следующим образом: нейросеть делает запрос огромного количества данных текстов, которые или предварительно были загружены в саму нейросеть, или находятся в интернете. Далее применяются специальные алгоритмы машинного обучения для того, чтобы нейросеть определила не только смысл написанного, но и классифицировала стили. «Технология, позволяющая получать от искусственного интеллекта тексты, называется Generative Pre-trained Transformer3, или GPT-3, которая переводится как генеративный предварительно обученный трансформер. Простыми словами, GPT-3 – это модель прогнозирования языка, в основе которой лежит глубокое обучение для создания текстов. Ее предыдущая версия могла обрабатывать 1,5 млрд параметров, нынешняя же – 175 млрд» [1].

Применение: генерирование текстов может применяться при написании статей, научных и исследовательских работ, различных обучающих заданий, литературных и стихотворных произведений (что было, в частности, продемонстрировано в рамках проведения форума «Россия» на ВДНХ в 2024 г.);

2) выполнение машинного перевода.

Для примера можно рассмотреть Яндекс Переводчик [5]. В базе Переводчика находятся миллионы переводов, а также огромное количество языковых моделей, которые система анализирует. Веб-служба начала переходить на нейронный машинный перевод с 2017 года, что качественно его (перевод) улучшило. «Согласно данным официального сайта Яндекса на 2021 год, Яндекс Переводчик использует гибридную систему (сочетает в себе статистический и нейронный машинный перевод) с помощью алгоритма CatBoost, учитывая недостатки обеих систем» [5]. В соответствии с актуальной информацией, в марте 2024 года перевод в Яндекс Переводчике доступен на 98 языков. Новшество нейронного перевода заключается в том, что идет не «дословный» перевод тек-

стов с языка А на язык Б, а точный перевод с высокой естественностью, учитывающий не только значение каждого отдельного слова, но и тональность текста и общий контекст.

Применение: быстрый, точный, качественный перевод текстов с высокой степенью естественности, написание исследовательских работ, обмен научными данными и достижениями, информацией, поиск и выжимка информации.

3) создание с помощью нейросетей различных чат-ботов.

Чат-боты являются настоящими виртуальными помощниками, а также искусственными собеседниками человека. Чат-бот позволяет с минимальными временными затратами получать нужную информацию, при этом заходить в браузер нет необходимости. Функционал чат-ботов постоянно расширяется [4]. Например, с помощью чат-ботов можно узнавать новости, скачивать книги, прослеживать актуальные курсы валют, узнавать прогноз погоды, рассчитывать стоимость товаров и/или услуг, быстро и точно получать обратную связь от различных организаций, включая медицинские клиники. С точки зрения лингвистического подхода, необходимо обратить особое внимание на текстовые и голосовые чат-боты, которые помогают в изучении и практике иностранного языка. Такие чат-боты создаются при помощи комбинации искусственного интеллекта, машинного обучения и технологии обработки естественного языка [4]. Чат-боты способны не только понимать запрос человека и давать обратную связь, но и обучаться новым командам, учитывая предыдущий опыт.

Применение: изучение и практика иностранного языка в области грамматики, лексики, стилистики. Практика письменного и устного перевода, аудирования, построения монологов и диалогов.

В заключении данной статьи можно сделать следующие выводы:

- 1) нейросети способны решать задачи высокого уровня сложности в различных сферах науки, бизнеса и образования;
- 2) технология обработки естественного языка (Natural Language Processing, NLP) дает возможность компьютерам интерпретировать человеческий язык;

3) нейросети могут использоваться в различных областях лингвистики, например, для написания научно-исследовательских работ, обучающих заданий, литературных и стихотворных произведений, выполнения быстрого и точного перевода на иностранные языки, обучения и практики иностранного языка.

Список литературы

1. Гуреева Ю. Топ-5 нейросетей для генерации текста / Ю. Гуреева [Электронный ресурс]. – Режим доступа: <https://rg.ru/2023/09/22/top-5-nejrosetej-dlia-generacii-teksta.html> (дата обращения 03.04.2024).

2. РБК Тренды. Что такое нейросеть: как устроен человеческий мозг «в цифре» [Электронный ресурс]. – Режим доступа: <https://trends.rbc.ru/trends/industry/641157be9a7947d3401fa3e8> (дата обращения 02.03.2024).

3. AWS. Что такое обработка естественного языка (NLP) [Электронный ресурс]. – Режим доступа: <https://aws.amazon.com/ru/what-is/nlp/> (дата обращения 01.03.2024).

4. Geek Brains. Что такое чат-бот: виды, способы, создания [Электронный ресурс]. – Режим доступа: <https://gb.ru/blog/chto-takoe-chat-bot/> (дата обращения 03.03.2024).

5. Wikipedia. Яндекс Переводчик [Электронный ресурс]. – Режим доступа: https://ru.wikipedia.org/wiki/Яндекс_Переводчик (дата обращения 03.03.2024).