

**Байкова Александра Васильевна**

д-р филол. наук, старший преподаватель

**Сунцова Александра Евгеньевна**

студентка

**Головина Анастасия Сергеевна**

студентка

**Овечкина Анастасия Владимировна**

студентка

ФГБОУ ВО «Вятский государственный университет»

г. Киров, Кировская область

## СПЕЦИФИКА НЕЙРОННОГО МАШИННОГО ПЕРЕВОДА

**Аннотация:** в статье анализируется вопрос эволюция машинного перевода от правило-ориентированных (RBMT) и статистических (SMT) систем к нейронному подходу (NMT), основанному на архитектурах глубоких нейронных сетей. Рассматриваются ключевые компоненты современных моделей – архитектура «кодер-декодер» и механизм внимания, обеспечивающие контекстуальную обработку языковых единиц и преодоление ограничений предшествующих парадигм. Особое внимание уделено роли параллельных корпусов как фундаментального ресурса для обучения моделей и инструмента эмпирического лингвистического анализа. Подчёркивается необходимость междисциплинарного взаимодействия лингвистов и специалистов в области искусственного интеллекта для дальнейшего развития технологий перевода.

**Ключевые слова:** нейронный машинный перевод, архитектура «кодер-декодер», параллельный корпус, обработка естественного языка.

Нейронный машинный перевод (neural machine translation, далее по тексту NMT) представляет собой современную парадигму автоматического перевода, базирующуюся на архитектурах глубоких нейронных сетей. До широкого внедрения нейросетевых подходов в области машинного перевода доминировали два

подхода: правило-ориентированный (rule-based machine translation, далее по тексту RBMT) и статистический машинный перевод (statistical machine translation, далее по тексту SMT). Правило-ориентированный подход базировался на ручной формализации лингвистических знаний: специалисты разрабатывали многоуровневые правила трансформации грамматических структур и лексических единиц между исходным и переводным языками, а также создавали обширные билингвальные словари. Несмотря на теоретическую прозрачность данной модели, её практическая реализация сопровождалась значительными ограничениями в переводе языковых явлений, что нередко приводило к снижению качества перевода в случае нестандартных или идиоматических конструкций.

Эволюционным этапом стало появление статистического машинного перевода, основанного на анализе параллельных текстовых корпусов. Данный метод автоматически извлекал вероятностные соответствия между языковыми единицами, опираясь на частотные закономерности их употребления в контексте. Хотя SMT обеспечил заметный прогресс в плане естественности перевода, он сохранял системные ограничения, а именно, искажение синтаксической структуры сложных предложений [2].

Необходимо отметить, что нейросети, применяемые в обработке естественного языка (далее по тексту NLP), включают разнообразные архитектуры, такие как рекуррентные нейронные сети (далее по тексту RNN), сверточные нейронные сети (далее по тексту CNN) и трансформеры. RNN представляют собой мощный класс архитектур, созданных для анализа последовательных данных, например текста, временных рядов и звуковых сигналов. Главное достоинство RNN заключается в наличии обратных связей, что дает возможность информации из предыдущих этапов влиять на текущие вычисления. Преимущество использования CNN для текстовых данных заключается в способности модели извлекать локальные и глобальные признаки из текста, что может улучшить способность модели к анализу и классификации текста. Трансформеры стали важным достижением в сфере обработки естественного языка и последовательной информации в общем. Одной из отличительных черт трансформеров является механизм внимания,

который позволяет модели концентрироваться на различных фрагментах входной информации в зависимости от их значения [1, с. 13, 33, 43, 140].

Одним из ключевых факторов, обусловивших широкое распространение нейросетевых моделей машинного перевода, стало существенное расширение доступности как масштабных билингвальных данных, так и вычислительных ресурсов. Параллельный корпус, являющийся одним из главных ресурсов для обучения моделей перевода, представляет собой структурированную совокупность текстовых пар на исходном и переводном языках, в которых единицы текста связаны отношением перевода. Критически важным этапом подготовки такого корпуса является процедура выравнивания, то есть установления соответствий между фрагментами текста на двух языках.

Рассмотрим пример из немецкого языка, иллюстрирующий сложность лексико-семантического соответствия между языками. Глаголы *beleidigen* и *kränken* в русскоязычной традиции лексикографии очень часто переводятся одним эквивалентом «обидеть», однако различия в их употреблении не всегда очевидны. Традиционные лексикографические источники, даже обладающие высокой полнотой охвата, не способны в полной мере передать подобные нюансы, поскольку они проявляются преимущественно в контекстуальном употреблении, а не в изолированном толковании. Именно в этой связи параллельные корпуса приобретают особую ценность как инструмент эмпирического лингвистического анализа. Позволяя исследовать реальные переводческие соответствия в естественных контекстах, они открывают доступ к скрытым паттернам употребления, недоступным при анализе словарных дефиниций.

Следует отметить, что процесс машинного перевода с применением нейросетевых моделей включает в себя ряд этапов, реализуемых в рамках архитектуры «кодер-декодер» с использованием механизма внимания (*attention*) [2]. Рассмотрим каждый из этапов подробнее.

1. Этап кодирования (*encoder*). На данном этапе входное предложение преобразуется в числовую форму, пригодную для обработки нейросетью: а) токенизация – исходное предложение разбивается на отдельные токены (слова или

подслова); б) векторизация – каждому токену сопоставляется вектор, отражающий его значение и контекстное окружение; в) контекстуальная обработка – кодер, как правило реализуемый с помощью рекуррентной нейронной сети (RNN), обрабатывает последовательность токенов, формируя их контекстуальные представления; г) формирование скрытого состояния – результатом работы кодера становится набор векторов, представляющих смысл всего предложения.

2. Механизм внимания (attention). Механизм внимания играет ключевую роль при генерации перевода: он позволяет модели «фокусироваться» на наиболее значимых частях входного предложения при формировании каждого слова перевода. На каждом шаге декодирования происходит взвешивание выходов кодера, в результате чего модель определяет, какие участки входной последовательности наиболее релевантны текущему контексту.

3. Этап декодирования (decoder). Декодер выполняет поэтапную генерацию перевода, используя информацию от кодера и механизма внимания: а) инициализация – процесс начинается с подачи специального стартового токена и начального скрытого состояния; б) генерация перевода – на каждом шаге декодер предсказывает следующее слово, опираясь на уже сгенерированные слова и контекст, полученный с помощью механизма внимания; в) завершение – генерация продолжается до появления специального токена окончания или достижения заданной длины последовательности.

Для иллюстрации работы системы рассмотрим перевод простого предложения.

1. Кодирование: предложение «Мой рейс был задержан» разбивается на токены. Эти токены преобразуются в векторы и обрабатываются кодером.

2. Механизм внимания: при генерации каждого слова на английском языке модель фокусируется на соответствующих токенах исходной фразы.

3. Декодирование: итоговый перевод формируется последовательно – “My flight was delayed”.

Таким образом, функционирование нейронных сетей в контексте машинного перевода представляет собой сложный многоэтапный процесс, реализуемый

в рамках архитектуры «кодер-декодер» с интеграцией механизма внимания. Нейронный машинный перевод следует рассматривать не как замену человеческому переводу, а как мощный инструмент, расширяющий возможности как прикладной переводческой практики, так и теоретической лингвистики. Его развитие требует междисциплинарного взаимодействия специалистов в области компьютерных наук, лингвистики и предметных областей, что подтверждает актуальность интеграции технологических и гуманитарных компетенций в современном образовании переводчиков.

### *Список литературы*

1. Картер Д. Нейросети. Обработка естественного языка / Д. Картер. – URL: <https://djvu.online/file/dDN01JlyxbP9t> (дата обращения: 15.02.2025).
2. Колганов Д.С. Обзор аналитической, статистической и нейронной технологий машинного перевода / Д.С. Колганов, Е.А. Данилов // Международный студенческий научный вестник. – 2018. – №3 (ч. 2). – С. 301–303. EDN YVQPWZ