

Баранов Иван Вячеславович

студент

Бояринцев Олег Андреевич

студент

ФГБОУ ВО «МИРЭА – Российский технологический университет»

г. Москва

СРАВНИТЕЛЬНЫЙ АНАЛИЗ ВЫБОРОК DEEPSEEK И GIGACHAT НА МАТЕРИАЛЕ ТЕХНИЧЕСКОГО ТЕКСТА ПО НЕЙРОСЕТЕВЫМ АРХИТЕКТУРАМ

***Аннотация:** в статье проводится сравнительный анализ двух нейросетевых моделей – DeepSeek и GigaChat – в задаче генерации учебного словаря технической лексики. Материалом исследования выступает научная статья по нейросетевым архитектурам. Обе модели получают идентичный промт с запросом на выделение слов, незнакомых русскоязычному студенту уровня В2, с переводом в контексте, транскрипцией и пояснением. Сравнение осуществляется по пяти критериям: количество найденных слов, точность перевода в контексте, сложность отработанной лексики, наличие пояснений и транскрипций, полнота охвата текста. В статье представлены результаты эксперимента для выборок объемом 25 и 50 слов, а также сформулированы рекомендации по использованию рассмотренных моделей в учебных целях.*

***Ключевые слова:** искусственный интеллект, GigaChat, DeepSeek, английский язык, сравнительный анализ.*

При изучении технического английского языка одной из ключевых проблем является корректный перевод терминологии в контексте. Лексические единицы, имеющие общеупотребительные значения в специальных текстах приобретают иной смысл. Стандартные словари и переводчики не всегда учитывают контекстные нюансы.

В данном исследовании проводится сравнение способностей двух нейросетевых моделей – российской GigaChat и китайской DeepSeek – генерировать

учебный словарь сложной технической лексики с переводом в контексте, транскрипцией и пояснением.

Материал исследования: научная статья «Some Attention is All You Need for Retrieval», посвященная анализу механизмов внимания в трансформерах и гибридных архитектурах SSM. Уровень сложности текста соответствует Graduate/Phd.

Промт единый для обеих моделей: Выпиши из текста [25/50] слов и фраз, которые могут быть незнакомы русскому студенту, изучающему английский на уровне B2. Для каждого слова дай: перевод на русский с учетом контекста этого конкретного текста, транскрипцию, краткое пояснение, почему это слово важно в данном контексте.

Критерии сравнения.

1. Количество найденных слов – соответствие фактического количества запрошенному.

2. Точность перевода контекста – учет технического значения термина в рамках данного текста.

3. Сложность найденных слов – доля лексики, объективно трудной для уровня B2.

4. Наличие пояснений и транскрипций – полнота дополнительной информации для каждой единицы.

5. Полнота охвата текста – равномерность распределения выбранных слов по структуре документа (начало, середина, конец).

Результаты тестирований.

DeepSeek, 25 слов

По критерию количества найденных слов модель продемонстрировала точное соответствие запросу: из 25 запрошенных единиц фактически получено 25. Оценка – 5/5.

Точность перевода в контексте составила 100%: ошибочных переводов не зафиксировано. Оценка – 5/5.

Анализ сложности отобранной лексики показал, что из 25 слов 16 относятся к уровням C1-C2 (сложные), 9 слов – к уровню B2 (средние). Доля сложных слов составила 64%, что соответствует оценке 3.5/5.

По наличию пояснений и транскрипций модель предоставила полную информацию: транскрипция и пояснение присутствуют для каждого из 25 слов. Оценка – 5/5.

Полнота охвата текста оценена как равномерная: слова распределены по всем разделам текста. Оценка – 5/5.

Итоговая оценка DeepSeek при запросе 25 слов составила 23.5/25.

Таблица 1

Результаты обработки DeepSeek 25 слов

Критерий	Оценка	Результат
Количество слов	5	25/25
Точность перевода в контексте	5	100%
Сложность найденных слов	3.5	64%
Наличие пояснений и транскрипций	5	Полное
Полнота охвата текста	5	Равномерный
Итог	23.5	23.5/25

GigaChat, 25 слов.

По критерию количества найденных слов модель продемонстрировала точное соответствие запросу: из 25 запрошенных единиц фактически получено 25. Оценка – 5/5.

Точность перевода в контексте составила 100%: ошибочных переводов не зафиксировано. Оценка – 5/5.

При анализе сложности отработанной лексики установлено, что из 25 слов 18 относятся к уровням C1-C2, 7 слов – к уровню B2. Доля сложных слов составила 72%, что соответствует оценке 4/5.

По наличию пояснений и транскрипций модель предоставила полную информацию: транскрипция и пояснение присутствуют для каждого из 25 слов. Оценка – 5/5.

Полнота охвата текста оценена как равномерная: слова распределены по всем разделам текста. Оценка – 5/5.

Итоговая оценка GigsChat при запросе 25 слов составила 24/25.

Таблица 2

Результаты обработки GigaChat 25 слов

Критерий	Оценка	Результат
Количество слов	5	25/25
Точность перевода в контексте	5	100%
Сложность найденных слов	4	72%
Наличие пояснений и транскрипций	5	Полное
Полнота охвата текста	5	Равномерный
Итог	24	24/25

DeepSeek, 50 слов.

По критерию количества найденных слов модель продемонстрировала точное соответствие запросу: из 50 запрошенных единиц фактически получено 50. Оценка – 5/5.

Точность перевода в контексте составила 100%: ошибочных переводов не зафиксировано. Оценка – 5/5.

Анализ сложности отобранной лексики показал, что из 50 слов 45 относятся к уровням C1-C2 (сложные), 5 слов – к уровню B2 (средние). Доля сложных слов составила 90%, что соответствует оценке 5/5.

По наличию пояснений и транскрипций модель предоставила полную информацию: транскрипция и пояснение присутствуют для каждого из 25 слов. Оценка – 5/5.

Полнота охвата текста оценена как равномерная: слова распределены по всем разделам текста. Оценка – 5/5.

Итоговая оценка DeepSeek при запросе 50 слов составила 25/25.

Таблица 3

Результаты обработки DeepSeek 50 слов

Критерий	Оценка	Результат
Количество слов	5	25/25

Точность перевода в контексте	5	100%
Сложность найденных слов	5	90%
Наличие пояснений и транскрипций	5	Полное
Полнота охвата текста	5	Равномерный
Итог	25	25/25

GigaChat, 50 слов.

По критерию количества найденных слов модель продемонстрировала точное соответствие запросу: из 50 запрошенных единиц фактически получено 50. Оценка – 5/5.

Точность перевода в контексте составила 100%: ошибочных переводов не зафиксировано. Оценка – 5/5.

При анализе сложности отработанной лексики установлено, что из 50 слов 45 относятся к уровням C1-C2, 5 слов – к уровню B2. Доля сложных слов составила 90%, что соответствует оценке 5/5.

По наличию пояснений и транскрипций модель предоставила полную информацию: транскрипция и пояснение присутствуют для каждого из 25 слов. Оценка – 5/5.

Полнота охвата текста оценена как равномерная: слова распределены по всем разделам текста. Оценка – 5/5.

Итоговая оценка GigaChat при запросе 50 слов составила 25/25.

Таблица 4

Результаты обработки GigaChat 50 слов

Критерий	Оценка	Результат
Количество слов	5	25/25
Точность перевода в контексте	5	100%
Сложность найденных слов	5	90%
Наличие пояснений и транскрипций	5	Полное
Полнота охвата текста	5	Равномерный
Итог	25	25/25

Полученные данные позволяют сделать следующие выводы. Обе модели успешно справились с задачей при обоих объемах выборки. При запросе 25 слов

GigaChat показал незначительное преимущество за счет более строгого отбора действительно сложной лексики. При запросе 50 слов обе модели достигли высоких результатов продемонстрировав 90% сложных слов и 100% точность перевода. Для задач составления словаря технического текста обе модели могут быть рекомендованы для использования. Выбор между ними может определяться личными предпочтениями.

Список литературы

1. Some Attention is All You Need for Retrieval. – URL: <https://ar5iv.labs.arxiv.org/html/2510.19861> (дата обращения: 19.05.2026).
2. Гусев А.К. Потенциал нейросетевых технологий для обучения английскому языку на ступени основного общего образования: вызовы и перспективы внедрения / А.К. Гусев, Т.Н. Озорнина // Уральский вестник образования. – 2025. – №4. – С. 23–30. – EDN HCYMXE.