

РАЗРАБОТКА СЕРВИСА ПЕРСОНАЛИЗИРОВАННОГО ОБУЧЕНИЯ НА ОСНОВЕ RAG С ГРАФОМ ЗНАНИЙ ПО УЧЕБНЫМ МАТЕРИАЛАМ

Шугаев М.Г.¹

Научный руководитель – канд. физ.-мат. наук Москаленко М.А.¹

¹Университет ИТМО

Введение

В современной образовательной среде наблюдается изменение паттернов поведения студентов: для решения учебных задач все чаще используются большие языковые модели (LLM), что позволяет быстро получать ответы, но снижает глубину погружения в предмет. Использование генеративного ИИ в этом аспекте может привести к получению поверхностных знаний, обучению на «галлюцинациях» языковой модели (недостовверных фактов), изучению материала, не соответствующего конкретной учебной программе [1]. С другой стороны, традиционный подход к изучению объемных первоисточников и учебников сталкивается с проблемами нехватки времени, отсутствием интерактивности и невозможностью оперативно получить разъяснения по сложным научным материалам. Следовательно, появляется необходимость в создании гибридного инструмента, который сочетал бы интерактивность LLM с достоверностью и строгостью оригинальных учебных материалов.

Основная часть

В работе предлагается архитектура сервиса, где LLM выступает в роли интеллектуального навигатора по загруженным пользователем учебным материалам. Техническая реализация базируется на усовершенствованном подходе Retrieval-Augmented Generation (RAG) [2]. Существующие нетривиальные решения («базовые» RAG-системы) имеют ряд весомых недостатков. Они демонстрируют низкую эффективность при работе со сложной технической документацией из-за потери структуры и разрыве смысловых связей при парсинге источников, не могут без специальной настройки качественно обрабатывать таблицы, картинки, формулы. Такие системы при диалоге с пользователем ориентируются на локальные совпадения из-за особенностей алгоритма векторного поиска релевантных материалов, не могут выявлять глобальные связи между разделами, терминами. Для решения этих проблем в работе предложены следующие методы

1) Интеллектуальный препроцессинг источников: вместо стандартных парсеров используется инструмент анализа верстки, позволяющий корректно обрабатывать многоколоночный текст, таблицы и формулы в формате LaTeX.

2) Обработка специальных объектов: для обработки изображений, таблиц и формул применяются визуальные языковые модели (VLM), генерирующие их текстовое описание для индексации и повышения релевантности выдачи ретривера.

3) Граф знаний: необходимое улучшение для образовательного сервиса диалогового формата, позволяющее выявлять глобальные связи в источниках [3].

Развитие функционала сервиса также предполагает возможность ранжирования материала на основе учебной программы, когда основные разделы подаются подробно, а дополнительные могут быть сжаты о краткого содержания с сохранением ссылок на оригинал. Такая возможность резонирует с проблемой дефицита времени у студентов. Завершающим этапом является модуль валидации знаний: система генерирует контрольные вопросы на основе пройденного материала, предотвращая формальное пролистывание текста.

Выводы

Разработанный прототип системы демонстрирует способность снижать уровень галлюцинаций большой языковой модели за счет опоры на пользовательские источники. Внедрение графа знаний и специализированной обработки материалов позволяет предоставлять студентам релевантные ответы с учетом контекста, основанном на полноценном понимании источника. Предложенный подход обеспечивает баланс между скоростью получения информации и глубиной ее усвоения, позволяя пользователям работать со сложными документами через современный интерфейс.

Литература

1. Edge D. et al. From Local to Global: A Graph RAG Approach to Query-Focused Summarization / Edge D. et al. [Электронный ресурс] // ArXiv.org : [сайт]. — URL: <https://arxiv.org/abs/2404.16130> (дата обращения: 18.02.2026).
2. Lewis P., Perez E., Piktus A. et al. Retrieval-Augmented Generation for Knowledge-Intensive NLP Tasks / Lewis P., Perez E., Piktus A. et al. [Электронный ресурс] // ArXiv.org : [сайт]. — URL: <https://arxiv.org/abs/2005.11401> (дата обращения: 18.02.2026).
3. Wang X., Wei J., Schuurmans D. et al. Survey of Hallucination in Natural Language Generation / Wang X., Wei J., Schuurmans D. et al. [Электронный ресурс] // ArXiv.org : [сайт]. — URL: <https://arxiv.org/abs/2202.03629> (дата обращения: 18.02.2026).

Шугаев М.Г. _____

Москаленко М.А. _____