

ИИ ЭКСПЕРТНАЯ СИСТЕМА – КОНСУЛЬТАНТ В СИСТЕМАХ СМЕШАННОГО ОБРАЗОВАНИЯ

Зарубов Е.Н.¹

Научный руководитель – Егошин А.В.¹

¹Университет ИТМО

353349@niuitmo.ru

Введение

Рост объёма цифрового учебного контента в системах удалённого образования усиливает проблему информационной перегрузки: обучающимся сложно быстро находить нужные материалы, которые распределены по различным фрагментам курса [1]. Для быстрого обращения к информации фактически требуется помнить структуру курса и расположение тем. При отсутствии этого знания обучающийся вынужден последовательно переходить между разделами и просматривать множество материалов, сопоставляя их содержание с формулировкой своего вопроса.

Чаще всего эту проблему пытаются решать структурированием материалов, навигацией, тегами и фильтрами, а также поиском и рекомендациями. Однако эти подходы требуют значимых трудозатрат на поддержание структуры и метаданных и остаются чувствительными к поиску по формулировкам запроса. На этом фоне большие языковые модели (LLM) рассматриваются как средство доступа к контенту через запросы на естественном языке и диалоговый формат взаимодействия [2–3]. Их использование в образовательных системах требует контроля достоверности и строгого ограничения ответа рамками курса, поскольку модели могут генерировать правдоподобные, но не подтверждённые материалами курса утверждения [4-5].

Основная часть

В рамках работы разработан LLM-ассистент, ориентированный на материалы LMS(Learning Management System), предназначенный для помощи обучающимся в поиске информации по материалам курса. Ассистент обеспечивает семантический отбор релевантных фрагментов по запросу пользователя, формирует краткую выжимку по найденным материалам и предоставляет ответ, формируемый строго по содержанию конкретного учебного курса.

Архитектура решения включает два сервиса: сервис на Kotlin принимает запросы из LMS, выполняет маршрутизацию и передаёт их в сервис обработки на Python, где реализован подход retrieval-augmented generation (RAG). В Python-сервисе выполняются предобработка и индексирование материалов курса, их хранение в векторной базе данных, извлечение релевантного контекста по запросу пользователя и взаимодействие с LLM для формирования итогового ответа на основе извлечённых фрагментов [6]. Такой подход повышает обоснованность результата и снижает риск неподтверждённых утверждений.

В ходе работы в LMS был реализован интерфейс диалогового обращения к материалам курса в формате чат-окна. На вход принимается запрос на естественном языке, на выходе формируется ответ по сути запроса и ссылки на соответствующие источники внутри курса, позволяющие перейти к первоисточнику и уточнить детали.

Ключевым ограничением является замкнутая предметная область: ассистент не использует внешние источники и не расширяет ответ за пределы доступных материалов курса. При отсутствии релевантных сведений или недостаточности контекста система формирует корректное сообщение об ограничениях.

Практическая применимость решения ориентирована на типовые запросы обучающихся: уточнение определений и формулировок и подбор фрагмента материала по конкретному пункту.

Выводы

В рамках работы разработан и внедрён LLM-ассистент для LMS, создаваемой в университете ИТМО на факультете ПИиКТ. Ассистент выполняет отбор релевантных учебных фрагментов по запросу пользователя, формирует краткую выжимку по найденным материалам и генерирует ответ, строго ограниченный содержанием конкретного учебного курса. При недостаточности релевантных данных в материалах курса предусмотрено формирование сообщения об ограничениях.

Перспективы развития включают расширение поддержки мультимодального контента (например, учёт транскриптов видео), а также инструменты для преподавателя по анализу типовых запросов и целевому улучшению структуры курса.

Литература

1. Ионов А. Ю., Ким Ю. Хен-Суновна. Цифровая усталость в реалиях современного образования // Современное педагогическое образование. 2025. № 6. С. 203–205. DOI: 10.24412/2587-8328-2025-6-203-205.
2. Бермус А. Г., Сизова Е. В. Педагогические, лингводидактические и психологические условия использования ChatGPT в системе высшего образования: систематический обзор // Концепт. 2024. № 11. С. 150–166. DOI: 10.24412/2304-120X-2024-11183.
3. Назаров Д. М., Бегичева С. В. Применение больших языковых моделей в образовательном процессе // Бизнес. Образование. Право. 2024. № 3(68). С. 430–436. DOI: 10.25683/VOLBI.2024.68.1057.
4. Айсин Т. Р., Шамардина Т. В. Детекция галлюцинаций на основе внутренних состояний больших языковых моделей // Электронные библиотеки. 2025. Т. 28. № 6. С. 1282–1305. DOI: 10.26907/1562-5419-2025-28-6-1282-1305.
5. Шевченко А. А. «Галлюцинации» ИИ как новая форма эпистемической ошибки // Respublica Literaria. 2025. Т. 6. № 4. С. 93–98. DOI: 10.47850/RL.2025.6.4.93-98.
6. Мельников А. В., Николаев И. Е., Русанов М. А., Аббазов В. Р. Сравнительный анализ методов RAG для построения русскоязычных интеллектуальных сервисов // Вестник Южно-Уральского государственного университета. Серия: Компьютерные технологии, управление, радиоэлектроника. 2025. Т. 25. № 2. С. 5–18. DOI: 10.14529/ctcr250201.