

Федеральное государственное автономное образовательное учреждение высшего
образования «Национальный исследовательский университет «Высшая школа
экономики»

Факультет компьютерных наук
Основная образовательная программа
Прикладная математика и информатика

КУРСОВАЯ РАБОТА
ПРОЕКТ НА ТЕМУ
"РАЗРАБОТКА ЧАТ-БОТА ДЛЯ АБИТУРИЕНТОВ"

Выполнил студент группы 244, 2 курса,
Толмаков Арсений Александрович

Руководитель КР:
Андреева Дарья Александровна

Москва 2026

1 Введение

Разработка чат-ботов для университета становится актуальной задачей в периоды пиковой нагрузки на службу поддержки университета. Большинство вопросов абитуриентов, студентов и остальных пользователей носит повторяющийся характер, и ответами на эти вопросы часто являются правила, требования и прочая информация, которая уже зафиксирована в нормативных документах НИУ ВШЭ.

Таким образом возникает возможность автоматизировать процесс взаимодействия со студентами и абитуриентами и снизить нагрузку на операторов службы поддержки: удобный чат-бот с понятным для обычного пользователя интерфейсом, и, что самое главное, с проверяемостью ответа, то есть пользователь будет получать ответ со ссылками на конкретный пункт.

Сам бот не будет являться полноценным собеседником – на него будут наложены ограничения для упрощения взаимодействия с пользователем, такие как: отсутствие диалогов на отвлеченные темы, отсутствие собственного мнения о сотрудниках университета и каких-либо мероприятий/событий, связанных с ним.

2 Цель и задачи работы

Цель работы

Разработать Telegram-бота для снижения нагрузки на службу поддержки НИУ ВШЭ в периоды повышенного спроса (периоды поступления и сессии). Бот будет представлять собой справочный сервис по документу ПОПАТКУС: поиск релевантной информации с указанием конкретного пункта и формирование пояснения по ней. В случае, если информация по запросу не найдена, бот будет давать контактную информацию службы поддержки.

Задачи работы

1. Подготовить заранее типовые запросы и сценарии
2. Спроектировать структуру работы бота: интерфейс, обработка запроса, модуль семантического поиска по документу и формирование ответа
3. Разработать механизм извлечения данных из ПОПАТКУС: чтение документа, его разбиение на фрагменты и их привязка к пунктам самого документа
4. Создать модуль LLM для разъяснения найденной информации и опорой в виде ссылки на источник
5. Добавить инструкцию для ошибочных запросов – разъяснение, предложения по исправлению и сообщение об отсутствии информации
6. Обеспечить фильтрацию ответа бота – отсутствие сбора личной информации пользователя, отсутствие диалога на отвлеченные темы, отсутствие какого-либо мнения о внутренних сотрудниках, соблюдение политик университета
7. Интеграция: связать Telegram-интерфейс, систему поиска информации по документу и генерацию ответа LLM в единую систему обработки запросов.
8. Тестирование работы бота: формирование 50 тестовых вопросов с заранее размеченными эталонными пунктами для расчета метрик
9. Добавить возможность оставить отзыв на ответ бота для будущих исправлений

3 Актуальность

Основным аргументом в пользу практической актуальности создания чат-бота в Telegram является снижение нагрузки на операторов службы поддержки во время набора новых студентов и периоды сессии. Автоматизация ответов на типовые вопросы значительно снизит время ответа на них и упростит работу службы поддержки.

Если говорить о технической актуальности используемого метода, очень важно учитывать, что ответ на вопрос пользователя должен быть как можно более строгим и при его нахождении его можно будет проверить. Поэтому в данном проекте будет использована архитектура RAG (Retrieval-augmented generation), которая объединяет в себе извлечение релевантных фрагментов из текста документа и формирование ответа языковой моделью на их основе. Такой подход отличается от простой генерации ответа языковой модели тем, что снижает риск возникновения галлюцинаций и додуманной информации. Также повышается надежность работы бота при обновлении документа: нет необходимости в переобучении модели. Это позволит быстрее поддерживать актуальность информации в документе.

Наконец, такой подход позволит проводить измеримое тестирование ответа и сформулировать критерии качества, что поможет улучшить работу бота.