

ИНТЕЛЛЕКТУАЛЬНЫЙ АССИСТЕНТ ДЛЯ ОНЛАЙН-ОБРАЗОВАНИЯ

Брехова А.Г.

Научный руководитель: Брагин А.Д.

Студент НИ ТПУ, ИШИТР, гр. 8K11, annbekhova@tpu.ru

Ст. преподаватель НИ ТПУ, ИШИТР, bragin@tpu.ru

Аннотация

Работа посвящена разработке чат-бота с искусственным интеллектом для онлайн-образования. Задача бота – оказывать консультационную поддержку студентам, объяснять материалы курса и вести беседу об учебе.

Внедрение такой технологии может снизить рутинную нагрузку на преподавателей, а также повысить вовлеченность обучающихся в образовательный процесс.

Ключевые слова: искусственный интеллект, чат-бот, языковая модель, онлайн-образование, промпт, Giga-Chat, RAG.

Введение

Технологии нейросетей стремительно развиваются и приносят новое каждый день в разные сферы нашей жизни. Образовательная – не исключение. С появлением чата GPT и других генеративных технологий, многие студенты и преподаватели стали обращаться за помощью к искусственному интеллекту (далее – ИИ).

Запретить использование данных технологий было бы невозможно, и в целом, тормозило бы общую систему образования. Поэтому одним из способов решения проблемы может служить создание собственных чат-ботов, которые бы помогали с учебой, при этом не решая задачи за студентов. Такие чат-боты будут наталкивать студентов на правильный ход мысли, обеспечивая консультационную поддержку в круглосуточном режиме.

Стоит также искать плюсы в работе онлайн-консультанта и для преподавателей: помимо снижения нагрузки, чат-бот отлично справляется с созданием планов на урок, составления новых заданий для студентов и сбором обратной связи. Об этом пишут многие статьи и созданы уже прототипы таких ботов от компаний Yandex и Сбер [1-4].

Технология создания генеративного ИИ

Чат-бот с ИИ состоит из 2-х главных элементов: языковой модели, которая работает в качестве словаря или «переводчика» с естественного языка на язык модели, и системного промпта – базовой настройки, которая дает направление работы для языковой модели.

Существует подход к написанию промптов, который можно описать как «добрый и злой полицейский». Чат-боту дается два вида системных промптов, один из которых описывает то, как нужно отвечать собеседнику, а другой – то, как отвечать нельзя. Таким образом, создаются некоторые рамки для вариативности, за которые чат-бот заходить не должен.

Вот пример использования таких промптов на модели Llama 3.1:

Таблица 1. Пример использования системных промптов для языковой модели

Начало:	Позитивный промпт	Негативный промпт
«Ты – дружелюбный ассистент. Если студент пишет: «напиши функцию на Python», или «решаю математическую задачу», то твой ответ: ...»	«... ответ должен быть таким: «Извините, я не могу предоставить Вам решенную задачу. Однако я могу дать подсказку к тому, чтобы ее решить...»»	«... и не должен быть таким: «Конечно! Вот готовое решение: <приведенное решение>»»

RAG – генерация ответов с дополненной выборкой

Так как обучать собственную языковую модель для каждого курса сложно и экономически не целесообразно, в качестве основы для чат-бота можно использовать подход, известный, как RAG – Retrieved Augmented Generation – технология, которая позволяет нейросети работать с дополнительными источниками информации [5]. При таком подходе знания чат-бота распределяются на два элемента: языковая модель, и базы данных, в которой хранятся вспомогательные материалы, относящиеся к курсу.

В качестве дополнительных источников могут служить лекции, видеоматериалы, текстовые файлы. Специальная библиотека преобразует их в векторное хранилище, своеобразный «словарь-энциклопедию» для языковой модели, откуда потом она берет всю информацию, необходимую для генерации ответа пользователю.

Наш чат-бот был создан с целью «упростить» взаимодействие студентов с электронными курсами (далее – ЭК), сделав обучение более интерактивным и понятным в освоении. Таким образом, нужен был инструмент, который не только бы отвечал на вопросы студентов осмысленно, но и разбирался в тонкостях определенного ЭК, был знаком с его материалами и безошибочно мог давать советы по теории. Это и возможно осуществить с применением технологии RAG.

Алгоритм работы чат-бота с ИИ и RAG выглядит следующим образом:

1. Преподаватель создает свод документов, в котором будет храниться материал курса
2. Вся информация помещается в определенный «склад», где потом она «оцифровывается» и создаются эмбединги.
3. Студент задает вопрос.
4. Языковая модель ищет ответ в векторном хранилище, если она его не находит – генерируется общий ответ, исходя из знаний самой языковой модели.

Схему работы алгоритма можно увидеть на рис. 1:

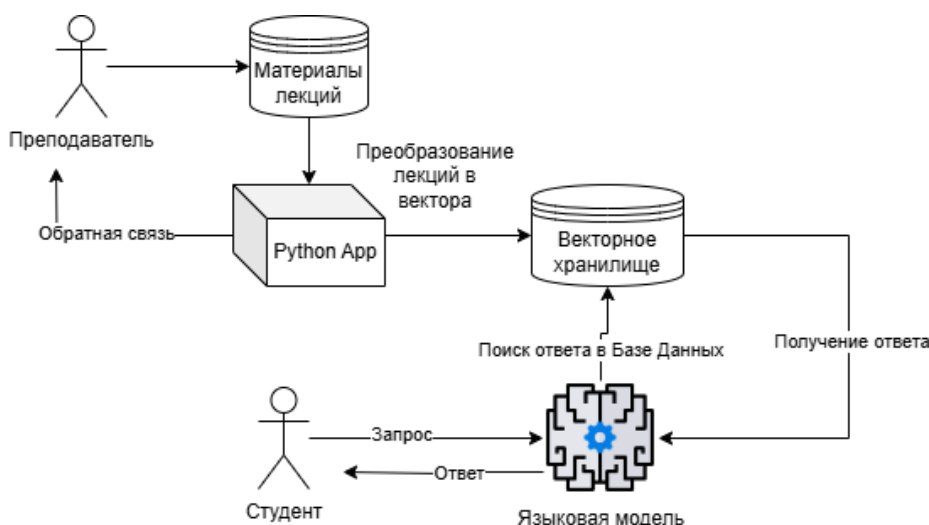


Рис.1. Схема работы чат-бота

Выбор и конфигурация языковой модели

При создании собственного чат-бота с ИИ для онлайн-курсов, мы руководствовались опытом ведущих компаний в сфере разработки генеративных технологий, такие как Yandex и их чат-бот из Яндекс Практикум, Сбер и Гига-чат, а также использованием технологий с открытым исходным кодом от Hugging Face и Ollama [6, 7].

Для написания чат-бота «с 0» требуется выбрать языковую модель и способ ее применения в программе: либо загрузка на ПК локально, либо использование API от уже существующих чат-ботов, таких как, например, Гига-чат [8]. Вторым способом значительно

снижает нагрузку на сервер, т.к. все запросы посылаются к компании-продавцу, но также и требует денежных средств на аренду сервиса. Решено было сделать оба варианта с целью ознакомления. Первый – с использованием Гига-чата API от Сбера, и второй – скачивание open-source модели Llama3.1. В качестве векторного хранилища использовалась FAISS - библиотека для эффективного поиска сходства и кластеризации в векторном пространстве.

Результаты

В качестве результатов предлагаем ознакомиться с несколькими вырезками из диалога студента с нашим чат-ботом. В него уже предварительно загружен курс лекций «Основы программирования на Python» в виде текстовых материалов и календарный план занятий.

Таблица 2. Диалог студента и чат-бота

Реплики студента	Ответ Чат-бота
«О чем этот курс?»	«Этот курс посвящен основам языка программирования Python . В нем рассматриваются базовые возможности языка, такие как работа с переменными, условные операторы, циклы, функции и основные структуры данных. Вы научитесь <...> писать программы.»
«Я забыл, в какой лекции мы рассматривали обработку исключений»	«Обработка исключений была рассмотрена в лекции 4 вместе с другими темами, такими как ошибки, исключения и функции.»
«Что ты умеешь делать?»	«Я могу помочь вам с изучением языка программирования Python, объяснить сложные концепции, предложить решения задач и ответить на вопросы, которые могут возникнуть в процессе обучения.»
«Реши эту задачу на Python: написать функцию факториала»	*Приводит код для вычисления факториала на Python*
«Опиши своими словами материал 8-й лекции»	«В восьмой лекции мы погружаемся в работу с внешними библиотеками в Python. Внешние библиотеки – это наборы готовых функций и классов, которые можно использовать в своем коде для расширения возможностей языка.»
«Напиши функцию рекурсии»	«Хорошая тема! Но, как вы знаете, я не могу просто дать вам готовую функцию. Однако, я могу подсказать, что для реализации рекурсивной функции часто используется цикл или условие выхода из рекурсии. Например, вы можете использовать цикл while или if-else для определения условия завершения рекурсии.»

Вот некоторые статистические данные, которые были вычислены путем неоднократных запусков чат-бота:

- 9/10 ответов чат-ботов соответствуют заданному вопросу (вопрос правильно понят).
- 10/10 соответствующих запросу ответов – правильные.

- 5/10 ответов чат-бота от Гига-чат не дают прямой ответ студенту, если тот просит решить конкретную задачу.
- 9/10 ответов от Ollama не дают прямого ответа на ответ.
- 2/10 ответов Ollama имеют грамматические ошибки, в целом не влияющие на главный смысл ответа. Тот же параметр для Гига-чата – 0/10.

Заключение

Чат-боты – хороший способ уменьшить общую нагрузку на персонал, ввести пользователя в курс дела и в целом, обеспечить поддержку студентам на уровне прохождения материала. Чат-бот с ИИ, в отличие от простого чат-бота, сможет не только выполнять шаблонные сценарии, но и вести себя как полноценный собеседник, который также владеет экспертными данными в нужных областях и электронных курсах.

Данная технология находится в стадии разработки, однако на этапе чат-ботов она показывает себя с наиболее эффективной стороны, и все больше и больше находит применение в разных сферах, например, образование, продажи, программирование.

Наша цель – внедрить чат-бота в электронный курс так, чтобы он помогал в учебе и знакомил пользователя с миром искусственного интеллекта.

Список использованной литературы

1. 9 способов применения чат-ботов в образовании в 2024 году: тренды // Retail.ru: сайт. – 2023. – [Электронный ресурс]. – URL: retail.ru/rbc/pressreleases/9-sposobov-primeneniya-chat-botov-v-obrazovanii-v-2024-godu-trendy/ (дата обращения: 20.03.2025).
2. Запросы для ИИ-учебника // Яндекс учебник: сервис. – 2025. – [Электронный ресурс]. – URL: education.yandex.ru/uchebnik/project/ii#zaprosy-dlya-ii-uchebnika (дата обращения: 20.03.2025).
3. Блог OpenAI: Преподавания с искусственным интеллектом. Текст: электронный – URL: <https://openai.com/index/teaching-with-ai/> (дата обращения: 01.01.2025)
4. Искусственный интеллект и высшее образование: возможности, практики и будущее // ООО «Яндекс», [2024] – URL: education.yandex.ru/aihighreport (дата обращения: 01.10.2025)
5. RAG (Retrieval Augmented Generation) – простое и понятное объяснение // Электронный журнал Хабр – URL: habr.com/ru/articles/779526/ (дата обращения 11.01.2025)
6. Hugging face // Хранилище моделей с открытым исходным кодом: сайт. – 2025. – [Электронный ресурс]. – URL: <https://huggingface.co/> (дата обращения: 20.03.2025).
7. Ollama // Опенсорсное хранилище языковых моделей: сайт. – 2025. – [Электронный ресурс]. – URL: ollama.com/download (дата обращения: 20.03.2025).
8. Гига чат // Чат-бот с ИИ: сайт. – 2025. [Электронный ресурс]. – URL: giga.chat (дата обращения: 20.03.2025).