

Кейс №2:

Разработка интеллектуального чат-бота «Я здесь живу» на основе технологии Large Language Model (LLM) с применением подхода Retrieval-Augmented Generation (RAG).

Краткое описание кейсового задания	Разработать интеллектуальный чат-бот "Я Здесь Живу" (ЯЗЖ) на основе технологии Large Language Model (LLM), способный отвечать на разнообразные вопросы жителей о повседневной жизни в Санкт-Петербурге и его районах. AI-помощник должен использовать открытые данные Санкт-Петербурга в качестве основы своей базы знаний, обрабатывая и интерпретируя эту информацию для предоставления актуальных и точных ответов на запросы пользователей, что сделает его полезным инструментом в повседневной жизни горожан и гостей города.
Полное описание кейсового задания	<p>Задача заключается в разработке интеллектуального чат-бота с использованием технологии Large Language Model (LLM). Этот AI-помощник должен быть способен отвечать на разнообразные вопросы жителей, касающиеся их повседневной жизни в Санкт-Петербурге и его районах.</p> <p>Ключевой особенностью чат-бота должно стать использование данных приложения «Я здесь живу» и других открытых данных Санкт-Петербурга.</p> <p>Приложение является частью экосистемы городских сервисов Санкт-Петербурга, развивается на платформе ВКонтакте и предлагает жителям 16 сервисов по различным жизненным ситуациям.</p> <p>Участникам будет предоставлен доступ к API приложения «Я здесь живу» и набор справочников, которые необходимо использовать при разработке решения, которое позволит боту предоставлять актуальную информацию по широкому спектру городских вопросов.</p> <p>AI-помощник должен уметь обрабатывать и интерпретировать эти данные для генерации релевантных ответов на запросы пользователей, делая его полезным инструментом в повседневной жизни горожан. Ориентируясь на категорию вопроса, предлагать подходящий сервис приложения «Я здесь живу» для поиска дополнительной информации.</p>

Проблематика	Особенностью задачи по разработке AI-помощника, является «очеловечивание» чат-бота «Я Здесь Живу»
Образ результата	<p>Разработать интеллектуальный чат-бот "Я Здесь Живу" (ЯЗЖ) на основе технологии Large Language Model (LLM) с применением подхода Retrieval-Augmented Generation (RAG).</p> <p>Использование LLM для генерации ответов.</p> <p>Реализация RAG для эффективного поиска и использования релевантной информации из базы знаний.</p> <p>Интеграция с API открытых данных Санкт-Петербурга для получения актуальной информации в реальном времени.</p> <p>Реализация системы логирования и анализа запросов для дальнейшего улучшения работы бота.</p> <p>Внедрение принципов обеспечения этичного и безопасного взаимодействия с пользователями</p>
Стек технологий, рекомендуемых/обязательных к использованию	Python, Docker
Ссылка на датасет	<p>Ссылка на датасет:</p> <p>Спецификация:</p> <p>а. Формат данных «csv»,</p>
Требования к железу	Нет ограничений
Уровень участников	Профи
Рекомендуемый состав команды для решения кейса	Специалист ML, специалист DS, специалист LLM, Промт-инженер

Теги на сайт (Тематика, Стек технологий)	AI-помощник, LLM, ML, Bigdata
Необходимые данные, дополнения/пояснения/уточнения <i>Информация для участников хакатона, которая поможет разобраться в проблематике и предложить качественные решения кейсового задания</i>	Для выполнения задания, предлагаем изучить действующий мини-апп «Я Здесь Живу» и его чат-бот.
Сделать кейсовое задание более привлекательным для участников	Информация не предоставляется

Методы и критерии оценивания

Для оценки решений применяется метод экспертных оценок, которые выставляют члены жюри, автоматизированные средства оценивания точности работы решений, и/или их сочетание.

На основании описанных ниже характеристик члены жюри выставляют оценки по шестибальной шкале 0–5 баллов.

Технический эксперт оценивает решение по следующим критериям:

- Запускаемость кода
- Обоснованность выбранного метода (описание подходов к решению, их обоснование и релевантность задаче)
- Релевантность результатов работы алгоритма
- Скорость формирования ответов
- Адаптивность/Масштабируемость
- Отсутствие в решении импортного ПО и библиотек, кроме свободно распространяемого с обоснованием выбора
- Наличие интеграционных интерфейсов, в первую очередь интерфейсов загрузки данных

Отраслевой эксперт и/или представитель постановщика задачи оценивает решение по следующим критериям:

- Релевантность поставленной задачи (команда погрузилась в отрасль, проблематику; предложенное решение соответствует поставленной задаче; проблема и решение структурированы)
- Уровень реализации (концепция/прототип и тд)
- Проработка пользовательских историй
- Реализация в решении требований Заказчика
- Выступление команды (умение презентовать результаты своей работы, строить логичный, понятный и интересный рассказ для презентации результатов своей работы)
- Читаемость слайдов
- Соблюдение единого стиля оформления
- Соблюдение принципов структурности
- Умение раскрыть тему
- Форма представления (творческий подход)
- Умение отвечать на вопросы (лаконичность и аргументированность)

Автоматизированные средства оценивания точности работы предложенных участниками алгоритмов (решений) выставляют оценку в диапазоне 0 - 1, где 1 равно 100% точности работы решения.

Итоговая оценка определяется как сумма баллов всех экспертов: технического, отраслевого и/или представителя постановщика задачи, как значение, выданное автоматизированными средствами оценивания, либо как итоговый балл жюри, умноженный на оценку автоматизированной системы.