

Поляков Егор

Мужчина, 24 года, родился 7 сентября 2001

+7 (967) 0117294 — предпочитаемый способ связи • TG: @Imuseless0
ep.rabota@mail.ru

Проживает: Москва

Гражданство: Россия, есть разрешение на работу: Россия

Не готов к переезду, готов к командировкам

Желаемая должность и зарплата

Junior Data Scientist

Специализации:

— Дата-сайентист

Занятость: полная занятость

График работы: полный день

Желательное время в пути до работы: не имеет значения

Опыт работы — 1 год 11 месяцев

Ноябрь 2023 —
настоящее время
1 год 11 месяцев

Справедливая Россия — За правду

Москва

Data Scientist

Описание проекта ИИ - агента:

Задача: AI - решение для анализа тона писем и последующей классификации о необходимости передачи документа главе партии.

Что было сделано:

- Реализован графовый пайплайн через LangGraph с архитектурой Planer - Thinker - Executer.
- Была развернута векторная база данных, на основе которой была сделана RAG - система с прецедентными случаями.
- Был добавлен инструмент для возможности поиска в интернете ИИ - агентом в случае недостатка информации в RAG - системе.
- Был реализован парсер ответов ИИ - агента с автоматической починкой парсера при некорректной работе одного.
- С помощью LangSmith произвел оценку качества работы модели и оптимизировал качество работы ИИ - агента с помощью промпт - инжиниринга
- Реализовано оптическое распознавание символов для обработки пользовательских документов.

Итог: данное решение позволило автоматизировать процесс принятия решений для автоматической пересылки документа главе партии и исключить человеческий фактор.

Описание проекта для транскрибации видео:

Задача: Разработать решение для транскрибации видео и аудио файлов и развернуть локально в пользовательской подсети.

Что было сделано:

- Была интегрирована модель OpenAI Whisper.
- Реализован API с несколькими эндпоинтами.
- Реализован интуитивный пользовательский интерфейс для работы с моделью.
- Произведено развертывание проекта через Docker Compose на сервере.

Итог: проект позволил оптимизировать работу отдела СМИ, так как им часто приходилось вручную транскрибировать видео и аудио файлы.

Образование

Высшее

2025	РТУ МИРЭА Прикладная математика, Интеллектуальный анализ данных
2023	РТУ МИРЭА Программная инженерия

Тесты, экзамены

2021	РТУ МИРЭА Cisco Networking Academy, Сетевые технологии
------	--

Навыки

Знание языков	Русский — Родной Английский — B2 — Средне-продвинутый
---------------	--

Навыки	Python ООП Git Классическое машинное обучение Визуализация данных Математический анализ Математическая статистика SQL Теория вероятностей Big Data Kubernetes CI/CD MLflow Apache Airflow Apache Spark Docker Clickhouse Numpy Optuna LLM LangChain Ollama RAG ChromoDB PyTorch Linux Bash pandas NLP PySpark
--------	--

Дополнительная информация

Обо мне	<p>Ссылка на мой GitHub - https://github.com/ChillLover</p> <p>Описание проекта ИИ - агента: Задача: AI - решение для анализа тона писем и последующей классификации о необходимости передачи документа главе партии. Что было сделано: - Реализован графовый пайплайн через LangGraph с архитектурой Planer - Thinker - Executer. - Была развернута векторная база данных, на основе которой была сделана RAG - система с прецедентными случаями. - Был добавлен инструмент для возможности поиска в интернете ИИ - агентом в случае недостатка информации в RAG - системе. - Был реализован парсер ответов ИИ - агента с автоматической починкой парсера при некорректной работе оного. - С помощью LangSmith произвел оценку качества работы модели и оптимизировал качество работы ИИ - агента с помощью промпт - инжиниринга - Реализовано оптическое распознавание символов для обработки пользовательских документов. Итог: данное решение позволило автоматизировать процесс принятия решений для автоматической пересылки документа главе партии и исключить человеческий фактор.</p> <p>Описание проекта для транскрибации видео: Задача: Разработать решение для транскрибации видео и аудио файлов и развернуть локально в пользовательской подсети. Что было сделано:</p>
---------	--

- Была интегрирована модель OpenAI Whisper.
- Реализован API с несколькими эндпоинтами.
- Реализован интуитивный пользовательский интерфейс для работы с моделью.
- Произведено развертывание проекта через Docker Compose на сервере.

Итог: проект позволил оптимизировать работу отдела СМИ, так как им часто приходилось вручную транскрибировать видео и аудио файлы.

Описание пет - проекта Obesity:

Задача: End to end разработка ML решения для классификации степени ожирения на основе данных о человеке.

Что было сделано:

- Оптимизация гиперпараметров моделей проводилась с помощью Optuna, что повысило значения метрик на 15%.
- Был развернут MLflow в облаке яндеса для логирования и трекинга результатов экспериментов.
- Реализован FastAPI - микросервис на основе лучшей модели и Gradio интерфейс для работы с микросервисом.
- Реализовал CI/CD пайплайн в Gitlab для автоматического деплоя в кластер Kubernetes.
- Реализовано автоматическое переобучение модели через некоторый промежуток времени с помощью Airflow.

- Знаю SQL.
- Знаю Git.
- Знаю принципы ООП.
- Основы Python (matplotlib/ seaborn, pandas, numpy, scipy, sklearn, catboost, optuna, mlflow).
- Знаю классические алгоритмы ML (лог. регрессия, лин. регрессия, деревья, случайные леса, градиентный бустинг, xgboost, catboost, lightgbm, svm, k-means, knn).
- Знаю статистику, теорию вероятности, мат. анализ и линейную алгебру.
- Знаю основы CI/CD.
- Знаю основы Kubernetes, Docker, Airflow.
- Создавал ИИ - агентов, RAG - систему для ИИ - агента, развертывал векторную базу данных для RAG - системы.
- Максимизировал качество работы созданного ИИ - агента путем промпт - инжиниринга.

Интересы:

end to end разработка моделей машинного обучения.

На должности младшего системного администратора обрабатывал наборы старых данных с разных информационных систем, приводя их в подходящий вид для переноса в новую информационную систему.