



Александр Ф. – Data Scientist, ML engineer

Ключевые навыки

FastAPI, ETL, Seaborn, MS SQL Server, Hadoop, Machine Learning, Python, Теория вероятностей, Математическая статистика, SQL, pandas, XGBoost, Scikit-learn, Matplotlib, PyTorch, Numpy, Git, TensorFlow, NLP, Clickhouse, PySpark, MLflow, Docker, PostgreSQL

Профессиональный опыт

RacoonSoft

Data Scientist

Август 2024 — Июль 2025 (11 месяцев)

Разрабатывал и внедрял систему прогнозирования изменения котировок на основе анализа текстовых данных и исторических данных о ценах, конкретнее:

- спроектировал и обучил несколько моделей LLM (включая BERT, GPT, RoBERTa, T5) для обработки (в том числе суммаризации) текстовых данных;
- разработал архитектуру для интеграции сигналов от моделей LLM с временными рядами, обученными в LSTM;
- использовал Scikit-learn, Numpy и Pandas для предварительной обработки данных и создания дополнительных фичей;
- использовал Docker для развёртывания окружений и запуска экспериментов с моделями на сервере;
- опыт работы с PostgreSQL: написание SQL-запросов для выборки и агрегации данных, создание и модификация таблиц, индексов и связей;
- настраивал и использовал MLflow для отслеживания экспериментов, логирования метрик и параметров моделей. Организовал хранение артефактов и версионирование моделей, что упростило воспроизводимость и дальнейший деплой;
- настраивал airflow и дашборды через streamlit.

Второй проект о создании модели генерации текстовых описаний столбцов в датасетах, включая:

- проектирование и обучение моделей: обучил модель LLM для генерации текстов на основе метаданных столбцов, учитывая терминологию и контекст;
- создание пайплайна обработки данных: реализовал этапы предобработки данных с использованием Pandas и Numpy для нормализации входных метаданных и генерации дополнительных признаков;
- разработку кастомной логики постобработки текстов: настроил библиотеку NLTK для улучшения синтаксической структуры генерируемых описаний, что повысило читаемость текстов;
- настройку инфраструктуры: автоматизировал процесс обновления данных и

перегенерации текстов через Airflow;

- использовал Docker для развёртывания окружений и запуска экспериментов с моделями на сервере;
- опыт работы с PostgreSQL: написание SQL-запросов для выборки и агрегации данных, создание и модификация таблиц, индексов и связей;
- настраивал и использовал MLflow для отслеживания экспериментов, логирования метрик и параметров моделей. Организовал хранение артефактов и версионирование моделей, что упростило воспроизводимость и дальнейший деплой;
- создание интерфейса визуализации: разработал веб-приложение на Streamlit для тестирования и демонстрации результатов генерации.

Стек: airflow, MLflow, streamlit, pytorch, sqlalchemy, transformers, nltk, pandas, Numpy, Scikit-learn, AWS S3, PySpark, PostgreSQL, Docker

Отраслевой центр разработок и внедрения информационных систем / ОЦРВ

Data Scientist, ML engineer

Август 2022 — Август 2024 (2 года 1месяц)

Делал чат-бота на основе NLP для автоматизации поддержки клиентов в компании.

- 1) Провел анализ исторических данных о взаимодействиях клиентов с поддержкой. (Pandas, NumPy, Matplotlib)
- 2) Занимался проектированием чат-бота: Разработал модель, способную понимать и обрабатывать запросы клиентов, используя NLP. Обучил модель на наборе данных диалогов с клиентами, чтобы бот мог распознавать намерения и отвечать на часто задаваемые вопросы. (PyTorch, NLTK, SpaCy, Transformers)
- 3) Интегрировал чат-бота с веб-платформой компании и CRM-системой для автоматического взаимодействия с пользователями. Обеспечил возможность передачи сложных запросов от чат-бота к живым операторам. (Flask, SQLAlchemy)

Итогом стало снижение нагрузки на службу поддержки на 30%, улучшение скорости ответа на запросы клиентов и повышение общей удовлетворенности пользователей. Регулярный мониторинг и анализ работы чат-бота позволили быстро выявлять и устранять недостатки, улучшая его эффективность и точность ответов.

Стек: Python, Pandas, NumPy, Matplotlib, NLTK, SpaCy, Transformers, PyTorch, Flask, SQLAlchemy, SQL

Образование

Высшее

Самарский национальный исследовательский университет имени академика С.П.

Королева, Самара

Фундаментальная информатика и информационные технологии