Станислав Королев — Data Scientist

Россия · tg: @lupinerh · lupinerh@gmail.com · Hugging Face

Обо мне

ML/NLP Data Scientist с опытом end-to-end разработки сервисов: от данных и архитектуры до деплоя и мониторинга. Запускал production-решения: FAQ-бот для автоматизации поддержки, модели LTV и churn для оптимизации маркетинговых бюджетов.

Опыт работы

Data Scientist — Independent Developer

2023 – Present

- RAG Movie Recommender Demo/HuggingFace: Создал рекомендательную систему с классическими подходами (SVD, Collaborative Filtering) и семантическим поиском Vector DB (ChromaDB/Faiss). Внедрил RAG-пайплайн, позволяющий подбирать фильмы по текстовому описанию с помощью LLM.
- Hybrid Sentiment Chatbot Demo/HuggingFace: Создал сервис анализа тональности текста (ML + BERT + LLM API). Реализовал визуализацию эмоциональной окраски на уровне слов.
- Educational AI Bot Demo/Telegram: Спроектировал и запустил сервис в Telegram для подготовки к собеседованиям. Динамическое формирование ответов на вопросы по DS/ML и систему повторения изученного материала с LLM API.
- Anki AI Assistant: Разработал веб-сервис для обучения с LLM. Реализовал функции автоматической оценки ответов, генерации тестов и аудио-подкастов.
- ToDo Assistant: веб-сервис управления задачами с LLM.
- Finance Planner: веб-сервис планирования бюджета с LLM.

Data Scientist — Plarium

05/2019 – 09/2022

- LTV & Churn Pipeline (ML): Построил автоматизированный пайплайн (Airflow) прогнозирования LTV. Модель позволила маркетингу оптимизировать долгосрочные бюджеты.
- Churn Prediction (ML): Разработал специализированную модель оттока для сегмента пользователей, повысив метрики удержания игроков.
- Revenue Forecasting (ML): Реализовал модель временных рядов для прогноза выручки от событий. Создал интерактивный инструмент симуляции сценариев.
- Fraud Detection System (ML): Реализовал модель для поиска нечестных игроков, выявив новые паттерны мошенничества.
- Support Automation Bot (NLP): Разработал E2E микросервис (API, Docker, MongoDB) для автоответов на вопросы пользователей. Решение (ML, NLP) превзошло штатный поиск и снизило нагрузку на техподдержку.
- Support Ticket Auto-Tagging (NLP): Ускорил обучение и инференс Transformer-модели для автотегирования тикетов.
- Optimal Push System: Создал алгоритм персонального рекомендованного времени push-уведомлений.
- User Segmentation: Провел кластеризацию пользователей по поведенческим паттернам и реализовал API для использования сервиса другими командами.
- Data Quality Monitor: Автоматизировал поиск аномалий в данных и дашборд в Tableau, сократив время реакции инженеров на сбои.

Образование

Магистр — Прикладная математика и информатика

Кубанский Государственный Университет (КубГУ), 2014 – 2020

Навыки

DS & ML: Python, SQL, PySpark, Pandas, NumPy, Scikit-learn, CatBoost, SHAP, LIME, Tableau

NLP: LLM, RAG, Transformers, BERT

MLOps: Docker, Apache Airflow, Git, REST API

Databases: PostgreSQL, MongoDB, Vector Database (ChromaDB), Hadoop