Тростонецкий Роман

ML/AI-инженер

Telegram @romantrost | +7 (915) 106-64-87 | trostobusiness@gmail.com | github.com/smurkiooo

Опыт работы

Агентство «САПФИР» (Сколково) | Стажировка июнь – октябрь 2026

- Разрабатывал e2e-систему мониторинга ключевых показателей развития ИИ в России. Строил пайплайны обработки данных и агентные сценарии на LangGraph, использовал LangSmith для трассировки и отладки.

Пет-проекты

RiskEmbed GitHub

- На 5,5 млн ежемесячных выписок AmEx обучил каузальный трансформер для построения эмбеддингов истории клиентов.
- Добавил эмбеддинги к агрегированным признакам в LightGBM. На отложенной выборке клиентов Gini вырос с 0,92295 до 0,92363 относительно LightGBM на одних агрегатах. Прирост составил +0,00068 (95%-й доверительный интервал по парному бутстрэпу от +0,00017 до +0,00115).

SmartPropertyScraper GitHub

- Собрал пайплайн объявлений Циана с MinIO и DVC. Модель CatBoost с CLIP-эмбеддингами фото и PCA предсказывала цену квартиры. На валидации $R^2 = 0{,}73$ (1 374 объявления в датасете).
- Сделал семантический поиск квартир на FastAPI (multilingual-e5, ChromaDB, локальная LLM) и Telegram-бота для оценки стоимости.

DocChat GitHub

- Собрал локальный RAG-чат по PDF с цитатами с указанием страниц и фильтром релевантности. Объединил E5 и BM25 через RRF, генерацию подключил через LM Studio.
- На синтетическом тесте гибридный поиск находил все эталонные фрагменты в топ-6 для каждого из 16 вопросов, на которые в документах были ответы.

Хакатоны

«АльфаБудущее» (Альфа-Банк) октябрь 2026

- В команде разработал умный кредитный кабинет предпринимателя. Занимался нейросетевой фабрикой признаков для скоринга на эмбеддингах клиентов и дообучал Qwen3.5-4B с помощью PEFT.

Technoforge AgroTECH (ЭкоНива, ВШЭ) 3-е место, осень 2025

- Проверял гипотезы, строил признаки и обучал модели градиентного бустинга для подбора рациона коров с учётом целевого жирнокислотного профиля молока.

Образование

Финансовый университет при Правительстве РФ 2024 – 2028
Прикладная математика и информатика, профиль «Прикладное машинное обучение»

Яндекс Практикум 2026
Курс «ML-инженер с опытом»

Карьерный акселератор BYTE (Финансовый университет, Газпромбанк) 2026
Курс машинного и глубокого обучения, диплом с отличием

Навыки

ML и данные Python, SQL, DuckDB, pandas, NumPy, scikit-learn, scikit-uplift, CatBoost, LightGBM, Optuna
DL и LLM PyTorch, Transformers, PEFT (LoRA/QLoRA), RAG, vLLM, экосистема LangChain
Инфраструктура FastAPI, Docker, Airflow, Dask, Hadoop, MLflow, DVC, Prometheus, Grafana, Git
Языки английский C1 (IELTS), немецкий B2 (Goethe-Zertifikat)