

Сергунцов Роман

+7 (910) 431-99-28 | Москва | r.serguntsov@yandex.ru | github.com/serguntsov | t.me/serguntsov

Образование

МТУСИ

Московский технический университет связи и информатики

Москва, Россия

2022 – 2026

Опыт

NLP Data Scientist

ПАО Сбербанк

июль 2025 – н.в

Москва, Россия

- Разработан гибридный RAG для анализа отчетов международных банков и аналитических агентств. Решение включает автоматизированный OCR-парсинг документов и retrieval-пайплайн. Достигнута высокая полнота поиска ($\text{recall}@10 = 0.98$); Оценка качества с использованием LLM-as-a-Judge подтверждает высокую точность извлечения релевантной информации.
- Разработано решение с использованием ансамбля GenAI/ML моделей для обработки новостного потока и оценки сентимента. Построенные модели позволяют в реальном времени строить оценки экономической активности (аналог PMI), инфляционных ожиданий и геополитических рисков. Полученные индексы используются для прогнозирования динамики цен акций, ключевой ставки и других финансовых активов.
- Разработан MVP - мультиагентная система предназначенная для автоматизации рутинных задач экономистов. Система построена по принципу manager-workers, включает компоненты deep-research, RAG. Решение позволяет разворачивать сценарии по заданной фабуле, что существенно ускоряет и упрощает работу аналитиков.

Data Analyst

ФГАНУ ЦИТус

июнь 2025 – июль 2025

Москва, Россия

- Разрабатывал и оптимизировал SQL-запросы и функции для работы с большими базами данных. Проектировал хранимые процедуры и функции для автоматизации бизнес-процессов и ускорения анализа данных, улучшая производительность существующих решений. Поддерживал базы данных, включая функции для агрегации и трансформации данных.

Проекты

AIStudyHelper | *LangGraph, PyTorch, Qdrant, Docker*

- RAG агент с базой знаний информации из учебного канала (посты, фотографии, голосовые сообщения, документы). Агент помогает одногруппникам быть в курсе всех изменений в учебном ритме.

DRUG_FEEDBACK_NLP | *Pandas, PyTorch, transformers, sklearn*

- Полученная модель может быть использована для автоматической оценки эффективности лекарств по текстовым отзывам, что позволяет фармацевтическим компаниям и регуляторам оперативно отслеживать обратную связь от пациентов и принимать решения на основе данных. ROC-AUC = 0.967, F1-score = 0.936

predict_customer_churn | *Pandas, sklearn, Optuna, SHAP, seaborn*

- Разработал и сравнил несколько ML-моделей для прогнозирования оттока клиентов банка, достигнув метрики ROC-AUC = 0.936 на лучшей модели Gradient Boosting после оптимизации гиперпараметров и feature engineering. Решение позволяет выявлять клиентов с высоким риском ухода для своевременного применения программ лояльности и снижения оттока.

Технические навыки

Programming/ML: Python; PyTorch, Transformers, LangGraph/LangChain, Pandas, NumPy, Scikit-learn, RAG, MCP, re, NLTK

Infra/Tools: Linux, Bash, Git, Docker, HDFS, PostgreSQL, Qdrant