

Герман Пермяков

Email: germaneper@yandex.ru

Телефон: +7 (993) 907-93-60

[GitHub: German2104](#)

Проживает: Москва

Дата рождения: 21 апреля 2005 (20 лет)

ЦЕЛЬ

Получение стажировки на позицию **Data Scientist** для разработки и валидации моделей машинного обучения (классификация, регрессия, рекомендации), проведения статистических экспериментов (А/В-тесты), анализа данных на Python/SQL и визуализации результатов с целью решения продуктовых задач и профессионального роста.

ОБРАЗОВАНИЕ

Московский Авиационный Институт (МАИ)

Факультет: Прикладная математика и информатика

2023 – н.в. (3 курс)

Дополнительное образование:

- Курс по математике для **Data Science** — [karpov.courses](#)
- Курс по машинному обучению — Евгений Соколов
- Курс по алгоритмам, 2024 — ШАД
- Курс по NLP (обработка естественного языка) — ШАД (открытый GitHub)

ПРАКТИКА

Практика в Сбер (Full-Stack разработка)

Лето 2023

- Разрабатывал и поддерживал веб-приложение с использованием **Java (SpringBoot)** и **JavaScript**.
- Создавал интерфейсы на основе HTML, CSS и **React**.
- Взаимодействовал с базами данных (**PostgreSQL**), писал SQL-запросы.
- Работал с Git для контроля версий и командной разработки.
- Использовал технологический стек Сбера.

ПРОЕКТЫ

RAG-система в ЛИС МАИ (гибридный поиск + генерация с цитатами)

2023–2024

- **Контекст:** пользователи обращались к большому корпусу PDF/DOC/веб-страниц, стандартный поиск давал нерелевантные ответы.
- **Ingestion:** конвертация PDF/DOC, извлечение текста, нормализация кодировок, обогащение метаданными, чанкинг с overlap.
- **Поиск:** (Milvus).
- эмбединги **Jina-AI v3** (мультиязычность, до 8192 токенов).
- **Генерация:** **GPT-4o** с обязательными ссылками на источники и guardrails (JSON-схемы) против галлюцинаций.
- **Прод:** **Docker**, CI/CD; наблюдаемость (ошибки, p-latency); офлайн-метрики качества $Recall@k$, $nDCG@k$.

- **Результат:** рост релевантности ответов; средняя задержка 1 с при top-k=10.

Проект DevUp (ML-разработка)

2025

- Геймифицированное образовательное приложение с интеграцией локальной LLM через OpenAI-совместимый API.
- Спроектировал и развернул локальный LLM-сервис на базе vLLM с OpenAI-совместимым API, интегрировал его с Telegram-ботом, реализовал обёртку на базе серверной библиотеки KServe и подготовил датасет для тестирования модели.

НАВЫКИ

- **Языки и БД:** Python, SQL, PostgreSQL.
- **ML/DS:** Scikit-learn, **PyTorch**, XGBoost, CatBoost, SciPy, NumPy, pandas, Matplotlib, Seaborn, визуализация данных, теория вероятностей, алгоритмы и структуры данных, A/B-тестирование, Deep Learning.
- **NLP/LLM:** NLP, NLTK, HuggingFace Transformers, LLM, **VLLM**, **SGLANG**.
- **Бэкенд и MLOps:** FastAPI, Docker, CI/CD, OpenCV.

ДОСТИЖЕНИЯ

- Разработка backend-сервиса для проекта веб-такси.
- Успешное прохождение курсов Карпова и Соколова по математике и машинному обучению.
- Участие в финале AI-хакатона: разработали мультимодальную RAG-систему для интеллектуального поиска по PDF-документам. Система автоматически извлекает текст и изображения из документов (PyMuPDF), кодирует их в единое семантическое пространство (Sentence-Transformers, CLIP), сохраняет в Milvus и предоставляет API для поиска (FastAPI). Поддерживается поиск по смыслу и визуальному контексту. Реализация на Python.

ДОПОЛНИТЕЛЬНО

- **Языки:** Русский (родной), Английский (A2).
- **Хобби:** разработка веб-приложений, изучение новых технологий.