

Лев Бреловский

Data Scientist | ML Engineer

+7 909 570 42 09 | brelovskylev@mail.ru | telegram: @LevBrell | github.com/Golden-dreamer | Петрозаводск
| готов к релокации

ML Engineer с опытом разработки production-систем для анализа документов. Специализируюсь на OCR, NLP и Document AI: строю end-to-end пайплайны от обучения моделей до **FastAPI**-сервисов и деплоя в **Docker**.

Опыт работы

май 2026 – наст. время: разработка GraphRAG-системы (см. «Проекты»)

ЛИГРЕС

Data Scientist

март 2024 – февраль 2026

Москва, удалённо

- Юридические документы с рукописным текстом обрабатывались вручную. Задача — автоматизировать обработку документов. Разработал production-систему автоматической обработки (OCR + NER + классификация) на **FastAPI + Docker + RabbitMQ + S3** — более 10 000 документов в сутки.
- Внедрил CV-препроцессинг для качества OCR: увеличил пул документов, обрабатываемых автоматически, за счёт лучшего извлечения данных со сканов низкого качества. **OpenCV, torchvision**. Эксперименты — **ClearML, Hydra**.
- Результат: 80% документов обрабатывались без участия человека (раньше — все вручную). Сократил время ручной проверки юристов на 2 часа в день, команда стала обрабатывать на 300 документов больше.

июль 2023 – февраль 2024: самостоятельное обучение ML/CV, pet-проект по сегментации, хакатон (см. «Проекты»)

Военный инновационный технополис «ЭРА»

Data Scientist

июль 2022 – июнь 2023

Анапа

- Внутренний прототип OCR для бумажных документов: классификация и извлечение полей на сканах низкого качества и ксерокопиях. **Tesseract, EasyOCR, OpenCV, Python**.
- Прототип развернул на **FastAPI + PostgreSQL** с мониторингом: низко-confidence документы уходили на ручную проверку, точность отслеживалась через **PostgreSQL**.

Навыки

ML/DL: PyTorch, torchvision, Hugging Face Transformers, PEFT/LoRA, Scikit-learn

NLP & CV: Tesseract, EasyOCR, spaCy, NLTK, OpenCV

MLOps & Infra: FastAPI, Docker, RabbitMQ, S3, vLLM, GitLab CI, ClearML

Прочее: Python (NumPy, Pandas, asyncio), PostgreSQL, Neo4j, Linux (Ubuntu/Debian), Nginx | English — B2

Проекты и достижения

Knowledge Graph QA System (GraphRAG)

2026

Построил GraphRAG-систему для ответов на вопросы по структурированным данным: распарсил данные в графовую БД **Neo4j** (212k узлов, 899k связей), создал веб-интерфейс для запросов на естественном языке — LLM генерирует Cypher-запрос и возвращает ответ с опорой на граф знаний. стек: **FastAPI, JavaScript, Docker, LLM**. github.com/Golden-dreamer/animeGraphRAG

Railway Segmentation

окт. 2023

Сегментация ж/д путей с помощью U-Net. Датасет: 89 размеченных изображений, расширенный аугментацией до 984. **PyTorch, OpenCV**. github.com/Golden-dreamer/railway-segmentation

Хакатон

2023

3-е место, Всероссийский хакатон — сегментация городских сцен.

Образование

Петрозаводский государственный университет

2018 – 2022

Бакалавр, Прикладная математика и информатика

Дипломная работа: прогнозирование цен на рынке недвижимости с помощью ML (регрессионные модели, feature engineering на данных объявлений).