

Дмитрий Балобин

NLP ML Engineer

Москва, Россия

hanaan@internet.ru
t.me/d0rj3228
huggingface.co/d0rj

Опыт работы

Яндекс <i>Senior ML Engineer, Алиса на Поиске</i>	Январь 2026 – Сейчас 5 мес.
Разработка, поддержка и обучение кастомных LLM моделей на нескольких GPU кластерах;	
Райффайзенбанк <i>Senior AI Engineer, CIB</i>	Сентябрь 2025 – Январь 2026 4 мес.
Разработка AI copilot и автономных агентов для онлайн трейдинга;	
2ГИС <i>Senior NLP Data Scientist</i>	Январь 2024 – Сентябрь 2025 1 год 8 мес.
- Внедрение LLM в UGC раздел приложения - саммаризация отзывов, контекстная саммаризация для онлайн поиска; - Проведение технических собеседований в свою и смежную команды;	
Сбер <i>ML Engineer, Команда NLP B2C</i>	Июнь 2022 – Декабрь 2023 1 год 7 мес.
- Разработка RLHF шага обучения для LLM GigaChat, проведение исследовательских экспериментов; - Создание пилотов внедрения GigaChat в B2C и внутренние продукты банка. Создание retrieval части для бизнес кейсов RAG с помощью GigaChat; - Доработка пайплайна обучения SFT GigaChat, адаптация внутренних моделей;	
Сбер <i>Стажёр Data Scientist, Команда NLP B2C</i>	Декабрь 2021 – Май 2022 6 мес.
- Речевая аналитика телефонных звонков (кластеризация, визуализация); - Улучшение пайплайна обучения модели intent recognition, адаптация его под внутренние требования качества;	

Образование

МФТИ ФПМИ <i>Аспирантура: Компьютерные науки и информатика</i>	2025 – 2028
МФТИ ФПМИ <i>Магистратура: ИВТ, Машинное обучение и анализ данных</i>	2023 – 2025
РТУ МИРЭА <i>Бакалавриат: Прикладная математика, Анализ данных</i>	2019 – 2023

Статьи

Балобин Д. Ю., Сомов О. Д. Улучшение методов детекции сгенерированных текстов. – 2025.

Используемые технологии

Инструменты: PyTorch, Transformers, DeepSpeed, langchain, PydanticAI, langfuse, k8s, FastAPI.