

Бородин Матвей Вячеславович

+7 916 715 9563 | salam121asd@gmail.com | t.me/Djatva | github.com/matveylogee

Образование

НИУ ВШЭ

Бакалавриат, Факультет компьютерных наук, Программная инженерия

Москва

2022 — 2026

Релевантный опыт

Группа алаймента Быстрых ответов Алисы на Поиске

Стажер ML-разработчик, Яндекс

Москва, Россия

Декабрь 2025 — Наст. время

- Принял участие в работе над **всеми частями пайплайна** генеративных ответов поиска: от майнинга данных, сборки пуллов, **sbs-замеров** до дистилляции планера, **GRPO** обучения генератора на новом формате данных и экспериментов с learn-пулами и гиперпараметрами.
- Дистиллировали планер: сделали более **экономичную** версию, снизили число **обращений к поиску** на 30% и убрали жёсткую структуру плана, сохранив качество **инфоконтекстов** (больше свободы генератору при том же уровне релевантности).
- Переобучили генератор под новый формат данных с использованием **CERL** и **GRPO**: добились более коротких и лаконичных ответов, уменьшили долю **fallback-генераций**, при сохранении метрик.
- Интегрировал новый **reward** фактчека в пайpline метрик, реализовал **GPU-whitelist** для **квантованных ревордов**, исправил проблемы версионирования и сборки пайплайна метрик.

Реализованные проекты

Бенчмарк для оценки эффекта Reversal Curse в LLM

Analytics Pretrain YandexGPT, ШАД

Москва, Россия

Ноябрь 2025 — Наст. время

- Дипломный проект в рамках ВКР **ФКН НИУ ВШЭ** совместно с **ШАД** и командой аналитики **Pretrain YandexGPT** на тему исследования эффекта **Reversal Curse** в моделях Яндекса.
- По итогам ресерча текущих продовых **фактобенчей YaGPT** базовые модели уже достаточно хорошо воспроизводят одноходные факты, поэтому проверка модели на скайлл обращать факты возможна только её **дообучением на синтетических фактах**. Поэтому решили рассмотреть смежную проблему — **мультихоп запросы**: получить по нескольким сложным условиям список ответов.
- Собрали пул шаблонных **мультихоп** запросов пяти уровней сложности на базе **Wikidata** и **BFCL v4: multi-hop** через обращение к их **API** и **sparql** запросы. Перефразировали и очеловечили шаблонные запросы через YaGPT, чтобы они были более приближены к реальным пользовательским.
- Запустили пайpline инференса ответов на моделях **Алиса на поиске**, **НейроАлиса** и **YandexGPT Reasoner**. Также планируется провести замеры на базовых моделях Яндекса без поиска и ризонинга, например, **YandexGPT 5.1 Pro**.
- Качество ответов моделей замеряли через **LLM-as-a-Judge** с кастомным промптом на выдачу грейда ответу в сравнении с **gold** ответом.

Дополнительный опыт

Победитель Всероссийской открытой олимпиады школьников по математике

A/B Week Школы Анализа Данных | Яндекс, ШАД

2020 — 2021

Июль 2025

Тренировки. ML-забег по Reinforcement Learning | Яндекс

Сентябрь 2025

Технические навыки

Языки: Python, C++, Swift, Kotlin, Java

Технологии: PyTorch, SQL, Nirvana (yandex apache airflow analog), CatBoost, LoRA, Git

Скиллы: Machine Learning, Deep Learning, Classic ML, NLP, LLM, RL (CERL, PPO, DPO и GRPO), RAG