Артур Бабкин

+7 (937) 579-42-07 | artur.babkin.2015@gmail.com | github.com/ArthurBabkin | tg: @theother_archeee
Русский — родной | Английский — C1

Образование

Университет Иннополис
Бакалавриат, Факультет компьютерных наук
- Специализация: анализ данных и искусственный интеллект. GPA: 4.5 / 5.0.

Публикации

EMNLP 2026 (Industry Track) · From Production Traffic to Post-Training: Building a Self-Hosted LLM That Covers the Corporate Request Mix · соавтор · ACL Anthology

Опыт работы

Research Engineer (LLM Post-Training & Alignment) Июль 2025 — настоящее время
АО "Тбанк" · R&D-команда фундаментальных моделей, T-Tech
- Работаю на всём цикле обучения LLM: генерация обучающих данных, SFT, обучение с подкреплением, распределённое обучение, разработка бенчмарков и оценка.
- Построил масштабные мультиагентные LLM-пайплайны для генерации RL-сред и обучающих данных: 4000+ RL-сред с верифицируемыми наградами (разнообразных, решаемых, сложных и однозначных) и 150 000+ SFT-примеров, отфильтрованных через rejection sampling и другие проверки качества.
- Улучшил наши модели в следовании инструкциям (MultiChallenge, IFEval, внутренние бенчмарки), а также в RAG, multi-turn и агентной банковской поддержке (Tau3 banking, внутренние бенчмарки): модели с <30B активных параметров вышли на уровень state-of-the-art моделей на отдельных внутренних задачах.
- Для этого использовал SFT на Megatron и RL: GRPO на verl с многокомпонентным reward modeling, спроектированным против reward hacking (например, награды за цитирование в структурированном выводе, заметно снизившие галлюцинации), и RL-среды с верифицируемыми наградами (RLVR).
- Построил, автоматизировал через DAG в Apache Airflow и поддерживал 15 внутренних и внешних бенчмарков; занимался мультиязычной оценкой и помогал адаптировать внешние бенчмарки на русский.

NLP инженер Ноябрь 2024 — Июль 2025
ООО "Круассан Студио" · Разработка AI-продуктов под ключ
Telegram-бот для финансового анализа компаний
- Руководил командой разработки из 3 человек (NLP, backend) при масштабировании от идеи до MVP: Telegram-бот на aiogram с бэкендом на FastAPI, интеграцией с международными биржами и котировками, хранением в PostgreSQL — инвесторы получили ROI 1000%+, подтверждённый чеком на инвестиции по повышенной оценке.
- Разработал на LangGraph 3 агентов для онлайн-анализа финансового состояния компаний: по метрикам и мультипликаторам (EPS, EBITDA, P/E, P/S, PEG и др.); по новостному фону — с использованием Perplexity DeepResearch, парсинга RSS-фидов и дедупликации новостей по порогу схожести Jaccard на эмбеддингах; по публичной отчётности (например, SEC filings).

Система проверки юридических договоров
- Запарсил и построил векторный индекс по 413K+ фрагментам законодательства РФ; сравнил 10+ моделей эмбеддингов по качеству, скорости и размерности — поиск стал в 6 раз быстрее, индекс в 2,7 раза компактнее.
- Перевёл индекс с IVFFlat на HNSW: throughput вырос в 10–15 раз (до 760 rps), latency 12 мс → 1–2 мс при recall 98% — поиск перестал быть узким местом пайплайна.

NLP инженер | Co-founder Октябрь 2023 — Январь 2025
ООО "Диплит" · Платформа с ИИ-репетитором по математике
- Презентовал идею и продукт стейкхолдерам простым, нетехническим языком — привлёк под идею $25 000 живыми деньгами (инвестиции и гранты) и 2 видеокарты NVIDIA A100 для использования 24/7.
- Собрал команду из 7 человек и руководил ей, в том числе разработкой: довёл продукт до MVP с 2000+ пользователей и статуса топ-1 инновации года в сфере образования в России по версии ВШЭ и Альфа-Банка 2024 года (skillbox.ru).
- Спроектировал мультиагентную систему из 7 агентов с оркестрацией через LangChain/LangGraph на бэкенде FastAPI: у каждого агента свой набор инструментов через tool calling (RAG, OCR, математические калькуляторы и верификаторы) под разные сценарии помощи пользователям (решение, проверка, совместное решение, теория и т. д.), хранение в PostgreSQL — ответ пользователю меньше чем за 45 секунд.
- Среднее время в чате выросло с 1 до 4,5 минут (в 4,5 раза), удержание на 30-й день (D30) — с 3% до 14% (в 4,7 раза).
- Собрал LLM-систему-решатель математических задач: промпт-инжиниринг и Tree of Thoughts с генерацией и валидацией гипотез в процессе решения, с доступом к RAG — +30% на открытых вариантах ЕГЭ по профильной математике 2024; дообучение SFT+DPO через LoRA/QLoRA дало +45%.