

Габдрахманов Рамиль

Data Scientist (RecSys / Search / LLM)

ramil2911912@gmail.com | t.me/ramil2911 | github.com/Ramil2911 | LinkedIn

О себе

Data Scientist в e-commerce (поиск и рекомендации). Веду полный цикл разработки ML-решений: от формализации бизнес-задачи и офлайн-оценки до A/B-тестов, внедрения и поддержки в production. Специализируюсь на candidate generation и ranking, а также на задачах de-biasing, diversification и cold-start.

Опыт работы

ecom.tech (ex samokat.tech)

Data Scientist

09.2024 – н.в.

Москва

- Развиваю production-контур поиска и рекомендаций в e-commerce: candidate generation → ranking → post-processing; формализую бизнес-приоритеты в логику ранжирования и систему офлайн-оценки.
- Улучшаю качество персонализированного ранжирования через LTR-подходы, семантические признаки и эмбединги; решаю cold-start и сильный дисбаланс данных, чтобы сохранять релевантность и устойчивость выдачи.
- Снижаю эффект popularity bias и повышаю разнообразие выдачи: de-biasing, диверсификация и business-правила, контролируя компромисс «точность ↔ разнообразие» без деградации пользовательского сценария.
- Веду полный цикл ML-разработки: офлайн-оценка и дизайн экспериментов, A/B, внедрение и поддержка в production; развиваю пайплайны на PySpark/Polars и MLOps-стек (Airflow, MLflow, Feast, Kubernetes).

Яндекс Практикум

Наставник курса программирования (part-time)

08.2023 – 06.2024

Москва

- Менторил студентов по C# и инженерным практикам, проводил разборы проектов и кода.
- Участвовал в разработке/улучшении учебных материалов и вебинаров.

Проекты и хакатоны (избранное)

- **YOLOv8, speed from video (Минприроды) – 2 место.** GitHub
- **Action recognition (РЖД), SlowFast/MViTv2 – 2 место.** GitHub
- **NLP (VK x HSE) – transformer, 3 место;**
- **Ads optimization (ЦП 2024) – RMSE: 8.7 → 5.4.** GitHub

Навыки

Языки: Python, SQL, Bash, C#

ML/DL: GBDT, линейные модели; нейросети (MLP/CNN/RNN); transformers/LLM: GPT, BERT, CLIP, etc; graph neural networks; metric learning; autoencoders

Search/Recsys: Candidate generation, reranking; Learning-to-Rank; BM25, TF-IDF; diversification, de-biasing, cold-start

Research → Production: experiment design, offline evaluation, A/B; ML pipelines, feature stores; model serving/monitoring; reproducibility

Data/Infra (working knowledge): Spark, Polars, Pandas; ClickHouse; Airflow, MLflow, Feast; Docker, Kubernetes; S3/MinIO; Git, Linux

Образование

- **НИУ МИЭТ**, Программная инженерия, бакалавриат 2022 – н.в.
- **Центральный университет**, Машинное обучение, ДПО + магистратура 2025 – н.в.
- **DLS (ФПМИ МФТИ)**, DL/CV 2023
- **SMILES (Skoltech)**, ML summer school 2024

Языки

Русский — родной

Английский — B2