Alexander Tereshin
Machine Learning Engineer

📱 +995 591 319 552 | @ alexandr.tereshin@outlook.com | 🔗 alex-tereshin | 🐙 alexander-tereshin | ✈ tealandr

Machine Learning Engineer specializing in **NLP**, **LLMs**, **Agentic Workflows**, and **RAG architectures**. Architected high-throughput inference pipelines and production AI systems in fintech and e-commerce. Combines deep NLP theory with robust **MLOps** engineering practices to build scalable, reliable AI solutions.

EXPERIENCE

Teaching and Research Assistant   Sep 2025 – Present
Higher School of Economics (HSE University)

- Mentor graduate students across **Basic and Advanced NLP** courses, responsible for curriculum design, grading practical assignments, and giving exams ranging from foundational **Transformers** to **Multi-Agent Systems**.

Machine Learning Engineer   Jun 2024 – Present
Raiffeisenbank

- Developing a corporate **Agentic AI Platform** leveraging **Qwen-2.5-Coder** for autonomous code generation within isolated Micro Sandboxes. Deployed **FastMCP** servers for Agentic RAG and implemented a **Zero-Touch QA agent**, using the **Ragas** framework for evaluation;
- Engineered async internal data labeling pipeline (**Litestar, FastStream, RabbitMQ**) scaling throughput **10x** (to 2K/day) with **92% accuracy** via **Qwen2.5 Few-Shot CoT**;
- Designed and deployed 4 production batch pipelines, orchestrating **PySpark** jobs on **HDFS** via **Airflow** to inference a **LightGBM** model (**81% F1**). Integrated full-stack observability and drift monitoring using **Evidently AI**, **Prometheus**, and **Grafana**;
- Developed a Hybrid RAG system (**Milvus, BM25 + BGE-M3**) automating 80% of basic HR policy responses, reducing query resolution time by 99% (hours to seconds) with **80 ms retrieval latency**;
- Won **1st Place** in company-wide hackathon engineering an agentic recruitment platform for candidate scoring and simulations via **OpenAI Function Calling**, utilizing **Docling + Structured Output** for resume parsing.

Data Scientist   Jun 2023 – Feb 2024
VseInstrumenti.ru

- Run Data Science for three core products on a **12M+ MAU** platform, executing **40+ online A/B tests** using advanced variance reduction (**Stratification, Delta Method**), directly driving a **$25K** monthly revenue increase;
- Engineered reusable ETL components (**Python, Airflow, Docker**) reducing ad-hoc data analysis time by **80%**. Established a reporting framework for all product teams;
- Developed and deployed time series forecasting models (**SARIMAX, CatBoost**), achieving an average **11% MAPE**. Mentored junior data analysts on the end-to-end Time Series Forecasting pipeline;

Data Scientist   Jan 2023 – Jun 2023
Sportmaster

- Engineered a production-ready ML classification pipeline (**Postgres**, **Pandas**, **scikit-learn**), automating the identification of high-risk warehouse safety patterns and improving proactive risk mitigation.

EDUCATION

Higher School of Economics (HSE University)   Sep 2023 – Jun 2025
Master's degree in Machine Learning and High Load Systems   *GPA: 8.29/10*

Saint Petersburg State University of Civil Aviation   Sep 2016 – Jun 2021
Bachelor's degree in Maintenance of Aircraft and Engines

SKILLS

Programming & ML: Python, PyTorch, Transformers, RAG, Agentic Workflows, LangChain, Vector Databases

MLOps & Orchestration: Apache Airflow, Kubernetes, Docker, CI/CD, MLFlow

Data & Databases: PySpark, SQL, Hadoop, Hive, MongoDB, AWS S3

Backend & Tools: FastAPI, Litestar, Kafka, RabbitMQ, Redis, Bash, Linux, Git, Pytest, R, LaTeX

Languages: English (B2), Russian (Native)