Ilia Ukhin

Applied AI Engineer | LLM Products & Evaluation | Team Leadership

Email: ukhin.ilya.alex@ya.ru | Telegram: @mechmabot_work | Phone:

Open to relocation to Europe.

ML engineer with experience building and deploying LLM solutions for banking, B2B sales and advertising. Leads a 3-person AI evaluation team at Tochka Bank. Hands-on background in LLM fine-tuning, agentic workflows and benchmarking, supported by a six-year degree in mechanics and mathematics from Moscow State University.

Professional Experience

Tochka Bank | Applied LLM Engineer and Team Lead | Jan 2026 - Present

- Lead a 3-person team building response-quality evaluation for a B2B support AI copilot, including an annotation and feedback analysis harness; achieved 95% precision in distinguishing good vs. poor responses (team leadership since Jul 2026).
- Trained a LoRA adapter for action planning on B2B support-specialist logs, achieving quality comparable to ChatGPT on internal task-specific prompts.
- Created training data from real support-action sequences for an agentic LLM assistant in a LangChain-based multi-agent workflow.

VK | Applied LLM Researcher | Aug 2025 - Jan 2026

- Designed evaluation benchmarks and fine-tuned an LLM for code generation in YQL, a proprietary SQL dialect, preparing it for a multi-agent pipeline.
- Led technical reading groups on state-of-the-art agentic NLP methods.

Yandex | Software Developer | Dec 2022 - Jul 2025

- Built and deployed production LLM pipelines for analyzing B2B sales-call transcripts; prototyped AI agents for classified-services products.
- Developed GPT-based ad-text generation (+4% quality) and ML product-offer ranking (+77% relevance) for an SMB advertising platform.

Huawei, Russian Research Institute | Engineering Research Roles | 2020 - 2022

Separate appointments.

- Implemented MipNeRF-360 in PyTorch and investigated NeRF-based 3D image enhancement and indoor-scene neural rendering (2022).
- Developed Python hyperparameter-optimization tooling and MATLAB deep-learning workflows (2021); studied numerical methods for linear systems (2020).

Supervision & Mentoring

OPN1 (Educational Project No. 1) | Research Supervisor

Scientifically and technically mentored three teams in a program where university and school students solve real business problems: VLA View Planner - experimental VLA-based robotic-arm planning for object scanning; AI furniture visualization - inserting furniture into room photos for a manufacturer, for planned website integration; school scheduling - a school timetable application.

Technical Skills

AI / ML: LLM fine-tuning and evaluation, benchmarking, agentic AI, RAG, PyTorch, JAX, LangChain, Langfuse, scikit-learn, XGBoost, CatBoost.

Programming / deployment: Python (advanced), SQL, C++ (intermediate), FastAPI, Git, Linux, PostgreSQL, MySQL, Airflow.

Education

Lomonosov Moscow State University | Sep 2018 - Jun 2024

Specialist Degree (6-year program), Mechanics Division, Faculty of Mechanics and Mathematics | GPA 4.6/5.0

Thesis: Physics-Informed Neural Networks for Nonlinear Solid Mechanics: The Galin Elastoplasticity Problem.

University Gymnasium of Moscow State University | Completed grades 10-11.

Achievements

Winner, Phystech Physics Olympiad (Level I competition), Grade 11 (secondary school).

Languages

Russian (native), English (C1), German (beginner, uncertified). IELTS eTRF available upon request.