Taya Yakovenko — AI/ML Engineer

taisiia.yakovenko@gmail.com · github.com/tyakovenko · linkedin.com/in/taisiia-yakovenko · portfolio-46729.web.app

PROFILE

AI/ML engineer who builds and evaluates LLM and agent systems end to end, from data pipeline to deployed product. I pair models with human judgment for work in high-stakes environments, with a double background in mathematics and computer science: the math for honest evaluation and statistical rigor, the CS for shipping real systems instead of notebooks.

SKILLS

Machine Learning: LLM fine-tuning (LoRA/QLoRA), model evaluation, PyTorch, scikit-learn, computer vision, anomaly detection, statistical modeling

AI-Assisted Development: Claude Code, agent orchestration, context engineering, skills & MCP setup, multi-agent workflows, LLM APIs (Claude, Gemini, Qwen)

Engineering & Infrastructure: Python, Go, Linux, Docker, Kubernetes, Git, CI/CD, REST APIs

Data & Cloud: SQL, PySpark, Pandas, MongoDB, Azure ML, HuggingFace

EXPERIENCE

ML Researcher (Computer Vision) — UT Health San Antonio, Klykov Lab · Remote · 02/2026–Present

- Building a zero-shot ML workflow to detect regions of interest in 3D cryo-fluorescence brain-slice volumes, cutting the manual expert annotation that bottlenecks cryo-EM targeting (no experiment-specific training data).
- Designed a four-level pipeline: zero-shot segmentation (Cellpose-SAM) → semantic scoring with a biomedical VLM (BioMedCLIP) → rule-based spatial plausibility → a napari curation tool logging every decision to machine-readable provenance.

Risk LLM Fine-Tuning (NDA) — PayPal · Master's thesis · Remote · 2025

- Built an end-to-end LLM fine-tuning workflow (data prep + prompt design, LoRA/QLoRA) for credit-risk assessment under strict confidentiality.
- Benchmarked fine-tuned LLMs against classical baselines (Random Forest, SVM, XGBoost) on F1, KS, and MCC with multi-seed stability checks.

Data Analyst (ML Performance & Reporting) — Empower Analytics · Worcester, US · 01/2025–03/2025

- Built a human-in-the-loop pricing model that suggested freight quotes from historical data; analyst approve / correct / reject decisions fed batched retraining. Ran the analysis on time-to-price for analyst alone vs. analyst + model.
- Stabilized the live performance dashboard (Google Sheets → Python Spark/Pandas pipeline with input validation).
- Built the training-data ETL as a single Kubernetes pod (MongoDB → Snakemake → S3).

Software Developer Intern — Hanover Insurance Group · Worcester, US · 05/2022–08/2022

- Built a Python/Pandas automation parsing Excel claims data into an internal-system format (~15× faster than manual entry); worked in an Agile team of 5.

SELECTED PROJECTS

MIRA — real-time anomaly detection (Go). Concurrent three-stage streaming pipeline, one goroutine per live subscription, lock-free hot path, REST API + live dashboard; dual statistical detectors merged into a high-confidence tier. · github.com/tyakovenko/wpi-mira

BlogAI — deployed LLM app. Multi-step drafting (capture → draft → human edit): Gradio on HuggingFace Spaces + Telegram capture bot + Notion integration. Live and daily-usable. · huggingface.co/spaces/tyakovenko/blogAI

The Automation Paradox — data study + live dashboard. Fetch → clean → analyze ETL over three merged public datasets; K-Means + regression; 6-tab Streamlit dashboard + Quarto report with annual GitHub Actions auto-refresh.

EDUCATION

Worcester Polytechnic Institute (WPI) — Worcester, US

- M.Sc., Artificial Intelligence — 2026 · B.Sc., Mathematics (Honors) — 2024 · B.Sc., Computer Science (Honors) — 2024

LANGUAGES

English — Fluent · Ukrainian — Native · Russian — Native · German — A1