

Anasse El Boudiri

ML / AI Engineer

Lausanne, Switzerland | +41 79 486 25 12 | elboudiri.anasse@gmail.com
B Permit | anasse-elboudiri | anassee15 | anasse.vercel.app

AI engineer with 1.5 years shipping LLM agents in production (**400k+ users**). Master's thesis on agentic evaluation and migration to open-weight models. Seeking ML / AI Engineer roles in Switzerland from **February 2027**.

WORK EXPERIENCE

Master Thesis, Agentic Evaluation & Open-Weight Migration · *Groupe Mutuel* 06/2026 – 02/2027

- Designed a multi-turn evaluation method for a multi-agent assistant: an **LLM-simulated policyholder** (personas, 4 languages) converses until its goal is met, then LLM judges score the full conversation.
- Built a pluggable evaluator library, from rule-based checks to LLM judges verifying each claim against policy documents, with per-use-case thresholds and pass@k / pass^k metrics, used to **catch regressions and gate feature rollouts in production**.
- Migrating to **open-weight models** (Qwen, Kimi, Mistral): benchmarking them against the proprietary baseline, fine-tuning with LoRA on anonymised traces, and sizing self-hosted inference in Switzerland.

AI Engineer Intern · *Groupe Mutuel* 08/2025 – 06/2026

- Deployed an agentic chatbot to **400k+ users**, avoiding **~18k calls** in 6 months (**≈ CHF 360k** estimated savings).
- Built a **translation agent** with language detection so the chatbot answers in French, German, Italian and English.
- Set up tracing and monitoring with Langfuse, speeding up debugging and performance analysis.
- Built Databricks ETL pipelines to analyze **200k+ conversations**, using BERTopic topic modeling to categorize user messages and guide improvements.

Software Engineer · *Scientific Visual SA* 06/2025 – 08/2025

- Achieved a **100x speedup** in 3D clustering by optimizing the algorithm in C++.
- Designed data processing workflows for a computer-vision ML pipeline, removing the need for a **5-hour re-scan** per material.

Machine Learning Research Assistant · *IDIAP Research Institute* 09/2024 – 01/2025

- Created a novel video dataset and evaluation metrics for hand-gaze detection.
- Trained a computer vision model to speed up annotation, establishing a baseline **F1-score of 0.76**.

PROJECTS

Vision Voyager · Multimodal Minecraft Agent *GPT-4.1, VLMs, Ollama*

- Built the vision pipeline feeding the bot's first-person view to a VLM in the curriculum and critic agents.
- Served open LLMs (Codestral, Mistral-Small) via Ollama on SCITAS and benchmarked them against GPT-4.1.

Educational AI Assistant · Open-Source LLM for STEM MCQA *SFT, DPO, QLoRA, RAG*

- Fine-tuned Qwen3-0.6B-Base on 24k STEM MCQs (**+4%** accuracy) and trained a DPO model on student preferences (**73%** reward accuracy).
- Halved model size with QLoRA 4-bit (**1.1** → **0.5 GB**) at equal accuracy; added RAG with a fine-tuned retriever.

EDUCATION

EPFL · Master of Science. Data Science (GPA 5.4/6) 2023 – 2027

Relevant coursework: Machine Learning, Deep Learning, NLP, Foundation Models, High Performance Computing

HES-SO, Neuchâtel · Bachelor of Science. Computer Science (GPA 5.6/6) 2020 – 2023

Relevant coursework: Data Structures & Algorithms, Databases, Computer Systems, Networks, Web Development

SKILLS

Programming: Python, C / C++, CUDA, MPI / OpenMP, SQL, JavaScript, CI/CD, MLOps
ML / LLM: PyTorch, scikit-learn, Hugging Face, LLM agents, RAG, SFT / DPO, QLoRA, LLM evaluation, Langfuse
Infrastructure: Linux, Git, Docker, Kubernetes, Azure, Databricks, HPC

LANGUAGES

French (Native) | **English** (C1) | **Arabic** (B1)

INTERESTS

Basketball | Football | Robotics