Abdelilah Younsi

AI/ML Engineer

Paris-Palaiseau abdelilah.younsi@polytechnique.edu +33775840612 AbdelilahYounsi WebSite

Skills

Programming & Big Data: Python, SQL, C++, Flask, FastAPI, API Development, MongoDB, PySpark, Elasticsearch, Apache Hadoop
ML/AI Libraries: Pandas, scikit-learn, PyTorch, XGBoost, Optuna, Unsloth, HG Transformers, CrewAI, LLama.cpp
MLOps: Snowflake, Databricks, AWS, Azure ML, MLflow, Kubeflow, Airflow, Docker, Git, CI/CD
ML Skills:: NLP, Computer Vision, Time Series, Multimodal RAG, Agentic AI, Hypothesis Testing, Predictive analytics
Soft Skills: Agile/Scrum Methodology, Analytical Thinking, Rigor, Project Leadership, Communication, Collaboration

Education

l'X Ecole Polytechnique, Paris-Palaiseau, France *M2 Data Science* *Sept 2024 – Sept 2025*

- **Relevant Coursework:** Computer Vision, Optimization for data science, Generative models, Graphs Machine Learning, Causal inference, Big Data, Cloud architectures...

EMINES, Mohammed VI Polytechnic University, Morocco *Master of Engineering in Industrial Management, Data Science Minor* *Sept 2021 – Sept 2024*

- **Relevant Coursework:** Probability Theory and Statistics, Intensive Data-science courses, Macroeconomics and Finance

Lycée CPGE Ibn Timiya, Marrakech, Morocco *Sept 2019 – June 2021*
Classes Préparatoires aux Grandes Écoles (CPGE)

Experience

IDEMIA PUBLIC SECURITY - *Iris Recognition* *Courbevoie, France*
Libraries and Skills: PyTorch, PyTorch Lightning, MLflow, Optuna, Accelerate, Autonomy *April 2025 – Oct 2025*

- Trained and optimized a **Vision Transformer** for iris identification, achieving a **0.26% FRR** and reducing comparison workload from **16 pairwise checks** to a **single** comparison, dramatically accelerating biometric matching.
- Built a production-grade **MLOps** pipeline with **Docker, MLflow, and AWS**, improving reliability, reproducibility, and shortening model delivery time into operational environments.

CLEVERLYTICS - *Phosphoric Acid Production Optimization* *Ben Guerir, Morocco*
Libraries and Skills: SQL, Pandas, PySpark, sklearn, XGBoost, Git, Docker, Project Leadership *Aug 2024 – Oct 2024*

- Led the phosphoric acid optimization project, combining insights from 30+ papers with rigorous analysis of a **million-row process dataset** to improve solution purity from **93% to 95%**, generating measurable cost savings for the client.
- Developed **Random Forest & XGBoost time-series models** to predict optimal H_2SO_4 flow rates for sulfate control, and presented findings to plant engineers for upcoming deployment.

CLEVERLYTICS - *Iris Recognition* *Ben Guerir, Morocco*
Libraries and Skills: PyTorch, OpenCV, C++, OpenIRIS, Project Leadership *April 2024 – Sept 2024*

- Led the end-to-end Iris recognition project and served as main client interface, delivering a **World-Coin** based pipeline that improved **accuracy by 60%** on the **Ubiris** dataset and presenting results to guide biometric system decisions.
- Designed a two-stage DL approach (Autoencoder + Siamese CNN) that cut UBIRIS FRR from **50%** to **20%**, supported by weekly client presentations and clear technical documentation highlighting business impact.

AI/ML Projects

Local Multimodal Agentic RAG System GitHub *Sept 2025*
Libraries and Skills: llama.cpp, Qwen, Voxtral, Milvus, Docker, Streamlit, CrewAI

- Built a fully local multimodal RAG system with llama.cpp (Qwen, Voxtral) for document and audio processing, using Milvus vector database for semantic search, deployed via Streamlit and Docker with no external API calls.

AI Agent-Powered Flight Search Application GitHub *June 2025*
Libraries and Skills: CrewAI, Selenium, Streamlit, BeautifulSoup, Gemini CLI, LLM

- Developed an AI-powered flight search tool using a CrewAI multi-agent system and Selenium for web scraping, enabling automatic flight retrieval, comparison, and recommendation from Kayak via a Streamlit interface.

Fast style transfer with instance normalization GitHub *May 2025*
Libraries and Skills: PyTorch, Git, Docker, Streamlit

- Developed of a real-time style transfer system using a CNN autoencoder, deployed via an interactive Streamlit interface and containerized with Docker for simplified multi-platform deployment.

Fine-tuning a Small Language Model for Summarization GitHub Hugging Face *April 2025*
Libraries and Skills: Unsloth, HG Transformers, PEFT (LoRA), TRL, llama.cpp, ollama

- Fine-tuned Qwen 2.5-0.5B-Instruct on CNN/DailyMail using LoRA and 4-bit quantization for news summarization (+11-28% ROUGE), with GGUF conversion for llama.cpp/Ollama deployment.