

# Abdelilah Younsi

Data Scientist/ML Engineer Polytechnicien à la recherche d'un CDI

📍 Paris-Palaiseau    ✉ abdelilah.younsi@polytechnique.edu    ☎ (+33) 775-84-06-12    🌐 AbdelilahYounsi

## Compétences

**Programmation & Big Data:** Python, SQL, C++, Flask, FastAPI, Développement d'API, MongoDB, PySpark, Elasticsearch, Apache Hadoop

**Bibliothèques ML/AI :** Pandas, scikit-learn, PyTorch, XGBoost, Optuna, Unsloth, HG Transformers, CrewAI, LangChain

**MLOps :** Snowflake, Databricks, AWS, Azure ML, MLflow, Kubeflow, Airflow, Docker, Gitlab, Kubernetes, CI/CD

**ML Skills :** NLP, Computer Vision, Time Series, RAG Multimodale, Agentic AI, Tests d'hypothèses, Tests A/B

**Soft Skills :** Méthode agile/scrum, Esprit analytique, Rigueur, Autonomie, Travail en équipe, Communication et présentation

## Formation

**École Polytechnique (l'X), Paris-Palaiseau, France** *M2 Data Science* *Sept. 2024 – Sept. 2025*

- **Cours principaux :** Apprentissage par renforcement, Computer Vision, Optimisation pour la data science, Modèles génératifs, Graphs Deep Learning, Inférence causale, Big Data, Architectures cloud...

**EMINES, Université Mohammed VI Polytechnique, Maroc** *Sept. 2021 – Sept. 2024*

*Diplôme d'ingénieur en Management Industriel, Mineure Data Science*

- **Cours principaux :** Théorie des probabilités et statistiques, Data science, Macroéconomie et finance, Stratégie d'entreprise

**Lycée CPGE Ibn Timiya, Marrakech, Maroc** *Sept. 2019 – Juin 2021*

*Classes Préparatoires aux Grandes Écoles (CPGE)*

## Expériences Professionnelles

**IDEMIA Public Security – Reconnaissance de l'Iris** *Courbevoie, France*

*Bibliothèques :* PyTorch, Optuna, Accelerate, MLflow, Git, Docker, AWS SageMaker *Avril 2025 – Oct 2025*

- Entraînement d'un modèle **Vision Transformer** pour l'identification par Iris, réduisant le temps de comparaison et atteignant des performances compétitives sur données **open-source**, avec **0,26 % de faux négatifs**.
- Intégration du modèle dans une pipeline **MLOps** avec **AWS**, incluant packaging, déploiement en endpoint REST et suivi expérimental avec **MLflow**.

**ARDIAN FRANCE - Prédiction temporelle des Cash Flow** *Paris, France*

*Bibliothèques :* Databricks, sklearn, Pandas, PySpark, PyTorch Forecasting, Séries temporelles *Jan 2025 – March 2025*

- Développement d'un framework hybride combinant un Temporal Fusion Transformer (TFT) et le modèle financier Takahashi-Alexander pour **prédire les cash flows de fonds de private equity**, capturant dynamiques temporelles et fondamentaux économiques.

**CLEVERLYTICS - Reconnaissance de l'Iris** *Ben Guerir, Morocco*

*Bibliothèques :* PyTorch, OpenCV, C++, OpenIRIS *Avril 2024 – Sept 2024*

- Développement d'un pipeline d'identification basé sur l'iris en utilisant le code **open source** d'OpenIris (Worldcoin), améliorant la précision sur des jeux de données publics de **60 %**.
- Réduction du **taux des faux négatifs** de **50 % à 20 %** sur le jeu de données public UBIRIS grâce à une approche en deux étapes: un autoencodeur pour l'encodage et un **Siamese CNN** pour la classification de paires d'iris.

## Projets en Computer Vision et GenAI

**Système RAG multimodal agentique** *Sept 2025*

*Bibliothèques et compétences :* CrewAI, Whisper, Gemini API, Milvus, Docker, Streamlit, RAG

- Développement d'un système RAG vocal basé sur CrewAI multi-agents, intégrant la transcription Whisper, l'API Gemini et la base vectorielle Milvus pour fournir des réponses intelligentes aux requêtes multimodales.

**Application de recherche de vols avec agentique** [GitHub](#) *Juin 2025*

*Bibliothèques et compétences :* CrewAI, Selenium, Streamlit, BeautifulSoup, Gemini CLI

- Conception d'un outil de recherche de vols basé sur des multi-agents CrewAI et Selenium pour le web scraping, permettant l'extraction automatique, la comparaison et la recommandation de vols à partir de Kayak via une interface Streamlit.

**Transfert de style rapide avec normalisation par instance** [GitHub](#) *Mai 2025*

*Bibliothèques et compétences :* PyTorch, Git

- Développement d'un système de transfert de style en temps réel avec un autoencodeur CNN pour générer des images stylisées, optimisé avec pertes perceptuelles et instance normalization pour réduire les artefacts de batch.

**Fine-tuning d'un petit modèle de langage pour la synthèse de texte** [GitHub](#) *Avril 2025*

*Bibliothèques et compétences :* Unsloth, Hugging Face Transformers, AutoAWQ, PEFT (LoRA)

- Fine-tuning du modèle Qwen 2.5-0.5B sur CNN/DailyMail avec LoRA et quantification 4 bits, batch optimisé pour génération de résumés et évaluation ROUGE sous contraintes Colab.