

# Ingénieur en Mathématiques Appliquées et Informatique

## Spécialisé en Intelligence Artificielle et Data Science

Ayoub  
EL HOUDRI

Paris, France

+33 7 83 22 67 03

ayoubelhoudri@gmail.com

y-aoub.github.io

github.com/y-aoub

linkedin.com/in/elhoudriay



## Expérience professionnelle

- Mai 2023 **Data Scientist (Stage)** - Median Technologies (Sophia-Antipolis, France).  
(6 mois)
- Implémentation d'un modèle génératif de type GAN pour la segmentation tridimensionnelle des poumons à partir de scans CT de patients.
  - Développement et déploiement d'une pipeline pour la détection et la localisation des tumeurs pulmonaires.
  - Classification des tumeurs et analyse de leur malignité à partir de leur structure.
- Frameworks et Outils** : PyTorch, MONAI, XGBoost, MLflow, OpenCV, Nilearn, 3D Slicer, Singularity, Docker.
- Juin 2022 **Data Scientist & Software Engineer (Stage)** - Karmen (Paris, France).  
(3 mois)
- Implémentation d'un modèle OCR pour l'extraction de données structurées en table à partir de documents financiers.
  - Développement et déploiement d'un système en cascade pour la détection de fraudes documentaires.
  - Mise en place et gestion d'une pipeline de données déployée comme API interne, orchestrée par des pratiques MLOps.
- Frameworks et Outils** : PyTorch, OpenCV, LayoutParser, Tesseract, Pandas, MLflow, FastAPI, Streamlit, Dataiku.
- Mai 2021 **Machine Learning Engineer (Stage)** - Digimind Labs (Berlin, Allemagne).  
(3 mois)
- Mise en place d'un générateur de données synthétiques en 3D pour l'augmentation des données d'entraînement et de test.
  - Fine-tuning du modèle Pixel2Mesh pour la reconstruction de forme à partir d'une image unique.
  - Déploiement et monitoring du modèle en production après son intégration dans une plateforme spécifique au client.
- Frameworks et Outils** : Tensorflow, PyOpenGL, OpenCV, TensorBoard, FastAPI, Pytest, GitLab Actions, Blender.

## Formation

- 2022 – 2023 **Master de Recherche** - CY Cergy Paris Université / ENSEA / Laboratoire ETIS (CNRS) - Intelligence Artificielle et Systèmes Complexes.
- 2020 – 2023 **Cursus Ingénieur Grande École** - CY Tech - Mathématiques Appliquées et Informatique - Spécialité : Intelligence Artificielle.
- 2018 – 2020 **Classes Préparatoires aux Grandes Écoles d'Ingénieur** - Lycée Jean Bart - Mathématiques, Physique et Sciences Industrielles (MPSI) - Spécialité : Mathématiques, Physique (MP).

## Projets

- Modélisation Neuronale de l'Hippocampe : Réseau de Neurones à Spike** - Neurosciences, Traitement du Signal.
- Modèle BERT pour l'Aide à la Décision en Investissement ESG** - NLP, Séries Temporelles.
- Plateforme SaaS pour l'Analyse Automatisée d'Offres d'Emploi** - Data Science, Software Engineering.
- Explication du Prix des Contrats Futures d'Électricité en Europe** - Statistiques, Machine Learning.
- Implémentation du Modèle MNet : Segmentation d'Images Rétiniennes** - Deep Learning, Traitement d'Image.
- RAG vs QLoRa Fine-tuning : LLaMa2 sur des Données Médicales** - LLM, Deep Learning.
- Modèle de Mix Marketing : Optimisation des Dépenses Marketing** - Statistiques, Machine Learning.
- Modèle de Markowitz : Optimisation de Portefeuille** - Modélisation Mathématique, Statistiques.

## Compétences informatiques

**Langages de Programmation** : Python (Pandas, GeoPandas, Matplotlib, Numpy, Scipy, PyTorch, Jax, Numba, Transformers, LangChain, PySpark, SpaCy, Scikit-learn, OpenCV, XGBoost, LightGBM, Streamlit, FastAPI, Scrapy, BeautifulSoup, Selenium, MLflow), R, Matlab, SQL, C, Linux Bash.

**Langages de Balisage** : HTML, Markdown, LaTeX.

**Logiciels et Outils** : Git, GitHub Actions, Docker, Airflow, Kafka, AWS (S3, EC2, ECS, RDS, Lambda), Dataiku.

## Langues

Français Langue maternelle.

Anglais Courant.

Arabe Bilingue.