

Oumayma IHRISSANE

Data Engineering Student | Seeking PFA Internship

oumaymaihrissane@gmail.com

[Portfolio : /oumaymaih](#)

[GitHub : /oumaymaih](#)

0659848602

Profile

4th-year engineering student specializing in Data Engineering at ENSA Al Hoceima, seeking a PFA internship in data-driven and AI-related domains. Passionate about data, artificial intelligence, and scalable systems, with hands-on experience gained through academic and professional projects.

Education

Engineering Studies in Data Engineering

National School of Applied Sciences (ENSA)

2022 – Present

Al Hoceima, Morocco

Technical Skills

Programming : Python , SQL, R

Data Engineering & Big Data : Spark, Flink, Hadoop, Airflow, PostgreSQL, MongoDB, Cassandra, MinIO

Machine Learning and AI : Scikit-learn, TensorFlow, Keras, LangChain, FAISS, NLP (spaCy, NLTK)

Tools & Platforms : Docker, Git, Streamlit, Power BI, Jupyter Notebook

Professional Experience

Caisse de Dépôt et de Gestion (CDG)

Rabat, Morocco

AI & Data Engineering Intern

August, 2025

- Designed and implemented a **RAG-based chatbot** (Proof of Concept) using LLMs, FAISS, and LangChain for domain-specific information retrieval from unstructured data
- Engineered data pipelines to collect, clean, and structure data from websites and PDFs
- Developed an interactive Streamlit interface enabling real-time query processing and user interaction

Projects

TaaSim | Real-Time Mobility & Big Data Platform

- Engineered scalable streaming pipelines with Kafka, Spark, and Flink for real-time mobility analytics
- Implemented ML-based demand forecasting and trip matching with zone remapping (Porto → Casablanca)

Brain Tumor Detection System | Deep Learning & Medical Imaging [link](#)

- Built a deep learning pipeline for brain tumor classification and segmentation from MRI scans
- Integrated LLM + Contextual RAG module to generate a medical report with findings and recommendation

Crop Yield Prediction | Machine Learning

- Engineered a leakage-free end-to-end ML pipeline for crop yield prediction
- Evaluated models using cross-validation, Selected Random Forest achieving strong performance $R^2=0.98$

Procurement Data Pipeline | Big Data & Analytics

- Designed a distributed data pipeline (HDFS, Trino, PostgreSQL) to simulate a procurement system
- Integrated orders and supplier data to compute net demand, generate supplier orders, ensuring scalability and reproducibility

Certifications

CS50 - Introduction to CS

[Harvard University](#) | edX

08/2023 - 12/2024 | [link](#)

Introduction to data analysis

[Young Researchers Forum](#)

10/03 - 16/03 , 2025 | [link](#)

Languages

Arabic | French | English

Spanish , Chinese : Beginner

Leadership & Impact

• DiGigirlz STEM Mentorship Program

9th édition

| [2026](#) |

• Founder – “ AlgoWave ” Programming Club

| [2023](#) |

• Ambassador – Future Leaders Academy, 14th

Edition

| [2022](#) |