

Hafsa El-Mahdi

Étudiante en Ingénierie des Données / Python • SQL • Big Data • Data Analysis

 ER-RICH

 hafsaelmahdi.in@gmail.com

 www.linkedin.com/in/hafsa-el-mahdi

Profil

Étudiante en deuxième année d'ingénierie des données à l'École Nationale des Sciences Appliquées d'Al Hoceima (ENSAH), passionnée par l'analyse de données, le Big Data et l'intelligence artificielle. Je possède de solides compétences en programmation (Python, SQL, Java) ainsi qu'en manipulation et analyse de données. À travers mes projets académiques et mon stage en data science, j'ai développé une forte capacité d'analyse, de résolution de problèmes et de conception de solutions basées sur les données. Je souhaite aujourd'hui approfondir mes compétences en ingénierie des données et contribuer à des projets innovants dans les domaines du Big Data et de l'intelligence artificielle.

Formation Académique

2025 – 2026	Engineering Degree: Data Engineering-second year <i>ENSAH - National School of Applied Sciences Al Hoceima</i>
2024 – 2025	Engineering Degree: Data Engineering-first year <i>ENSAH - National School of Applied Sciences Al Hoceima</i> ✉
2022 – 2024	Preparatory Cycle: Science & Engineering Techniques <i>ENSAH - National School of Applied Sciences Al Hoceima</i> ✉
2021 – 2022	High School Diploma: Physical Science <i>Imam Ali High School</i>

Compétences

Programming Languages: SQL,PL/SQL, Python, R, MATLAB, C,JavaScript, PHP,Java (Spring, Spring Boot)	Bases de données : MySQL, PostgreSQL, MongoDB,Cassandra (NoSQL)	Visualisation des données: Power BI
web scraping BeautifulSoup, Scrapy	Développement Web : HTML, CSS	Big Data Spark, PySpark, Hadoop,kafka,Data Warehouse ,Data Lake
Analyse et traitement de données andas, NumPy, Data Cleaning, Data Analysis, Data Visualization	Intelligence Artificielle Machine Learning (scikit-learn, etc.), Deep Learning (TensorFlow, Keras, PyTorch) , NLP RAG	Outils : Docker,Apache Airflow , Git, Linux, Jupyter Notebook
Compétences douces: Gestion de projet,Leadership,Esprit d'équipe ,L'autonomie.	Bureautique : Microsoft office(word,Powerpoint,Excel)	Development Software Microsoft Visual Studio, IntelliJ,Dev++.

Expérience Professionnelle

Stage en Data Science

- Projets réalisés :
 - Prédiction de notes de films .
 - Classification des fleurs d'Iris .

- Prédiction des ventes .

désigner

- Club DatAi – ENSAH

Stage en AI

- Développement d'un Chatbot d'assistance universitaire
- Technologies : Python, NLP, RAG, LLM, Ollama, Data Processing
- Conception et développement d'un chatbot intelligent basé sur l'architecture Retrieval-Augmented Generation .
- Utilisation de modèles de langage locaux via Ollama pour générer des réponses aux utilisateurs.

Participation à la compétition internationale Hult Prize 2022

Participation à la compétition internationale d'entrepreneuriat social Hult Prize 2022, où notre équipe a développé un projet innovant visant à répondre à un défi sociétal mondial. Notre projet a été sélectionné pour être présenté lors de la phase internationale à Taipei, ce qui nous a permis de présenter notre solution devant un jury et de renforcer nos compétences en innovation, travail d'équipe, gestion de projet et communication.

Projets

Pipeline Big Data - Supermarket Supplies (Architecture Lambda)

- Technologies : Docker, Hadoop, HDFS, PostgreSQL, Trino/Presto, Airflow, Streamlit
- Automatiser l'approvisionnement des supermarchés avec ingestion, stockage, traitement fédéré et dashboarding

Pipeline de web scraping

- Développement d'un pipeline automatisé en **Python** pour l'extraction, le nettoyage et la visualisation des données.

Développement d'une application de gestion des congés en C

- Développement d'un système de gestion des congés en C, avec stockage et traitement des données des employés à l'aide de fichiers et structures. Conception d'une interface en mode console interactive permettant la gestion des demandes, la validation des congés et la gestion des entrées utilisateurs.

Pipeline de données distribuées avec MongoDB Sharding & Replication

- Technologies : MongoDB, Python, NoSQL
- Mise en place d'une architecture NoSQL distribuée permettant la scalabilité horizontale grâce au sharding.
- Configuration de clusters MongoDB avec replication pour garantir la disponibilité et la résilience des données.
- Gestion des datasets volumineux avec des stratégies de partitionnement adaptées.
- Monitoring et optimisation des performances des clusters MongoDB pour assurer une réponse rapide aux requêtes.

Langues

- | | |
|----------------------------|---------------------------|
| • Arabe : native | • Anglais : Intermédiaire |
| • Français : Intermédiaire | • Espagnol : débutante |

Centres d'Intérêt

— Peinture | voyages | Lecture | Guitare | Gaming | cuisine