



# Youssef CHAABI

Aspirant Data Engineer en recherche d'un stage de fin d'études de 6 mois

🏠 202 Rue de Vaugirard Paris 75015

✉ youssefchaabi55@gmail.com

📅 21 ans

☎ 06 22 15 50 11

📍 France

## Compétences

### Bases de Données

MySQL, OracleDB, MongoDB, Casandra

### ETL

Apache NiFi, Talend, Apache Airflow

### Big Data

Apache Spark, Apache Kafka, Flink

### Data Warehousing

Amazon Redshift, Snowflake

### Conteneurisation & Orchestration

Docker, Kubernetes, CI/CD

### Méthodologies et Gestion de Projets

Méthode Scrum, Modélisation UML, Merise

### Gestion de version et environnement

Git, GitHub, Unix/Linux

## Langues

### Français

Bilingue

### Anglais

Professionnel

## Centres d'intérêt

### Jeux d'échecs

### Mathématiques récréatives

## Réseaux sociaux

in @youssef-chaabi

🐦 @yChaaby

## Certification

**Microsoft Certified** Power BI Data Analyst Associate (en cours)

**AWS Certified** Cloud Practitioner

**IBM AI Engineering**

## Diplômes et Formations

### Master 2 : Exploration Informatique des Données & Décisionnel

Depuis septembre 2023

Université Sorbonne Paris Nord - Institut Galilée Île-De-France

### Licence en Sciences et Techniques : Génie Informatique

De septembre 2022 à juin 2023

Université Abdelmalek Essaadi - Faculté des Sciences et Techniques Tanger

### Diplôme d'Études Universitaires en Sciences et Techniques : Mathématique Physique Informatique

De septembre 2020 à juin 2022

Université Abdelmalek Essaadi - Faculté des Sciences et Techniques Tanger

## Expériences professionnelles

### Data Scientist - Stagiaire

D'avril 2023 à juin 2023

Institut National de Recherche Halieutique Casablanca, Maroc

- Automatisation de la classification des segments de pêche
- Développement, test et optimisation des modèles de classification, atteignant une précision de plus de **94%**
- Développement d'une application web permettant aux utilisateurs de faire des prédictions
- Collaboration au sein d'une équipe de production, gestion du projet selon la méthodologie **Scrum**

**Stack technologique** : Python, Apache Spark, PCA, Arbre de décision, Forêt aléatoire, XGBoost, LightGBM, CATBoost, Pandas, NumPy, Django, MVT

## Projets

### Chatbot - Fine-tuning d'un modèle Llama3-8B (en phase de conception)

- Fine-tuning du modèle Llama3-8B pour personnaliser les réponses aux besoins des futurs étudiants
- Collecte de données à partir des documents internes de l'Université

**Stack technologique** : Llama3-8B, LangChain, Solon-embeddings, PyTorch, MongoDB, Cassandra, MERN Stack

### Intégration et analyse des données crypto

- Automatisation de l'ingestion des données crypto provenant de diverses APIs, mise en place d'un Data Lake pour stocker les données brutes
- Visualisations interactives des tendances des crypto-monnaies

**Stack technologique** : AWS S3, Snowflake, Apache Airflow, Tableau BI, Python

### Plateforme de formations avec recommandation collaborative

- Streaming en temps réel pour la mise à jour dynamique des données.
- Automatisation de la mise à jour du modèle de recommandation.
- Optimisation continue du système en fonction des nouvelles données utilisateurs

**Stack technologique** : OracleDB, Spark, Apache Kafka, Docker, AWS Lambda, ALS (Alternating Least Squares), Python, Streamlit, Selenium