



Omar KHEMMI

Étudiant en Data Science

☎ +212 7 71 76 48 81 ✉ khemmiomar1@gmail.com

🌐 omarkhemmi 📄 Omar KHEMMI

Étudiant en dernière année de Master en Informatique et Data Science, passionné par la transformation des données en informations exploitables. Compétences en **apprentissage automatique**, **prétraitement des données** et **visualisation des données**. Expérimenté en **Python**, **SQL** et **Power BI**. Maîtrise de **Hadoop** et **MongoDB** pour l'analyse de Big Data. **Recherche un stage de fin d'études (PFE)** pour appliquer des solutions basées sur les données et approfondir ses compétences en **Intelligence Artificielle** et **Data Engineering**.

FORMATION

Mathématiques et Informatique pour la Data Science , Master	2024 - 2026
<i>ENSA Khouribga, Université Sultan Moulay Slimane</i>	Khouribga
Informatique Décisionnelle , Licence Professionnelle	2023 - 2024
<i>École Supérieure de Technologie (EST), Université Mohammed Premier</i>	Oujda
Génie Informatique , Diplôme Universitaire de Technologie (DUT)	2021 - 2023
<i>École Supérieure de Technologie (EST), Université Cadi Ayyad</i>	Safi

EXPÉRIENCE

Ingénieur Logiciel - IA NLP Ensa – Khouribga (2 mois)	Juil 2025 – Aoû 2025
<ul style="list-style-type: none">Création d'une pipeline complète pour l'entraînement de Small Language Models (SLM) sur des jeux de données spécifiques.Application de LoRA / PEFT et quantification pour optimiser les performances et comparaison avec de grands LLM.Livraison d'un SLM prêt pour la production avec scripts, documentation et rapports de performance.	
Développeur Full Stack ADS – Rabat (3 mois)	Avr 2024 – Juin 2024
<ul style="list-style-type: none">Développement d'une plateforme digitale pour l'intégration sociale des femmes avec visualisations BI interactives et pipeline ETL.Développement front-end avec Next.js et back-end avec Node.js, assurant un flux de données fluide et interactif.Intégration de tableaux de bord Power BI pour des insights en temps réel, fournissant une solution évolutive pour la prise de décision basée sur les données.	
Développeur Web Full Stack Devosoft – Safi (2 mois)	Avr 2023 – Mai 2023
<ul style="list-style-type: none">Conception et mise en œuvre d'une plateforme d'annuaire web utilisant Laravel, HTML, CSS, et JavaScript, améliorant l'accessibilité et l'organisation du contenu.	

PROJETS

WordCount Distribué sur Hadoop (Python Streaming)

- Implémentation d'un programme **WordCount MapReduce** avec **Python Streaming** sur **Hadoop**.
- Conception de scripts mapper et reducer pour traiter de grands ensembles de texte distribués.
- Expérimentations sur **Cloudera VM** pour analyser la scalabilité et explorer les concepts **Big Data** via des TPs sur **Pig**, **HDFS** et **MapReduce**.

Système d'Extraction de Connaissances d'Articles Scientifiques

- Développement d'un **système de data science** pour collecter et analyser les publications scientifiques (**arXiv**, **PubMed**, **Semantic Scholar**).
- Utilisation de **NLP** pour extraire entités et relations entre auteurs, institutions et concepts.
- Création d'un **tableau de bord interactif (Streamlit, Neo4j)** pour visualiser graphes de connaissances et tendances.

Plateforme de Prédiction de Matches de Football

- Création d'un **système de machine learning** pour prédire les résultats des matches de football à partir des données pré-match et de performance des équipes.
- Comparaison et optimisation des modèles **Régression Logistique**, **Random Forest** et **XGBoost** via validation croisée.
- Conception d'une **interface web interactive** pour la prédiction en temps réel et la visualisation des performances.

TECHNOLOGIES

Langages Core : Python, Java, JavaScript, SQL, POO

Data Science ML : Pandas, NumPy, Scikit-learn, TensorFlow, Matplotlib, Seaborn, XGBoost

Data Engineering Big Data : Power BI, Hadoop, Spark, Airflow, Pipelines ETL, Entrepôt de données (SQL Server)

Bases de données Stockage : MySQL, PostgreSQL, MongoDB, Firebase, Neo4j