

# Lachhab Fatima

☎ 07-72-87-28-49   ✉ [lachhatfati2020@gmail.com](mailto:lachhatfati2020@gmail.com)   📍 Casablanca   🔗 [LinkedIn](#)   🐙 [GitHub](#)   🌐 [Portfolio](#)

---

## PROFIL

Étudiant en master de science des données avec une solide formation en informatique et des compétences avérées en résolution de problèmes. Passionné par l'analyse des données et l'adaptation à de nouveaux défis.

## ÉDUCATION

**2023-2025 :** (En cour) Master en Science des Données et Big Data, Faculté des Sciences Ben M'Sik, Casablanca

**2020-2023 :** Licence en Mathématiques et Informatique, Faculté des Sciences Ben M'Sik, Casablanca

## PROJECTS

- **Time-Series Energy Consumption Forecasting**

Développement d'un modèle de prévision de la consommation d'énergie horaire avec XGBoost, atteignant un RMSE de 3 753,6 MW. En nettoyant plus de 145 000 points de données et en intégrant des facteurs comme les jours fériés et week-ends, nous avons amélioré la précision des prévisions. Ce modèle optimise l'allocation des ressources, réduit les coûts et assure la fiabilité du réseau électrique.

- **Prédiction de dons de sang avec régression logistique et SHAP**

Développement d'un modèle de régression logistique pour prédire les dons de sang futurs, en analysant les données avec Seaborn et Matplotlib. Le modèle, évalué avec une validation croisée, a atteint un AUC moyen de 0,7468. SHAP a été utilisé pour expliquer l'impact des différentes caractéristiques, optimisant ainsi les campagnes de collecte de sang.

- **MultiPDF Chat App**

Application en Python utilisant LangChain, FAISS, et des modèles de langage comme GPT-3.5 et HuggingFace. Elle permet de poser des questions en langage naturel sur des PDF et d'obtenir des réponses précises en moins de 2 secondes. Développée avec Streamlit, elle offre une solution rapide et intuitive pour l'analyse documentaire.

- **Lexical Disambiguation in Arabic Using AI**

Création d'un modèle basé sur BERT pour désambiguïser les mots arabes en analysant leur contexte via la similarité cosinus. Le projet a atteint une précision de 87 %, améliorant la traduction automatique et les systèmes de dialogue.

- **Breast Cancer Diagnosis Predictor**

Application Streamlit déployée sur Streamlit Cloud utilisant la régression logistique pour prédire si une masse mammaire est bénigne ou maligne avec une précision de 98 %. Les utilisateurs peuvent entrer des données cliniques et visualiser les résultats via des graphiques interactifs.

- **Blog Generator App**

Une application Streamlit exploitant Llama 2 et des prompts dynamiques pour générer des blogs personnalisés. Elle permet de créer des articles sur des sujets variés en quelques secondes, avec un contenu structuré adapté à différents styles d'écriture, automatisant jusqu'à 80 % du processus de création.

- **Rider Shop Interactive Dashboard**

Construction d'un tableau de bord interactif Power BI en transformant des données brutes de location de vélos avec SQL pour dégager des informations stratégiques. Il met en avant des indicateurs clés tels que 10M\$ de revenus, 7,03M\$ de bénéfices, et 2M d'utilisateurs, segmentés en 81,81% d'utilisateurs enregistrés. L'analyse des tendances horaires et saisonnières a permis d'optimiser la prise de décision et d'améliorer les stratégies marketing et l'efficacité opérationnelle.

- **Prédiction des prix des voitures**

Création d'un modèle de régression linéaire pour prédire les prix des voitures en fonction de plusieurs facteurs comme la marque, le modèle et le kilométrage. Après un prétraitement minutieux des données, j'ai évalué le modèle avec des métriques de performance comme le MSE et le  $R^2$ . Ce modèle permet d'obtenir des estimations précises des prix des véhicules, facilitant les transactions.

## COMPETENCES

- **Programmation :** C, C++, C#, Java, Python (Django, Numpy, Pandas, Scikit-learn), HTML, CSS, JavaScript.
- **Bases de données :** SQL (MySQL, Oracle), PL/SQL, SQLite.
- **Big Data :** Hadoop, Hive, Spark, Kafka, Elasticsearch, Kibana.
- **Systèmes d'exploitation :** Windows, Linux (Ubuntu), Shell Scripting.

## LANGUES

- Français (Courant)
- Anglais (Niveau professionnel)
- Arabe (Courant)