

# LEBNAITI Rajaa

Future Ingénieure en Data Science et IA | Spécialisée en Machine Learning et IA | A la recherche d'un stage de fin d'études (PFE)

rajaa.lebnaiti@gmail.com | +216 06743707 | Portfolio.com | linkedin.com/LebnaitiRajaa  
github.com/LebnaitiRajaa

## Profil

Étudiante en dernière année d'Ingénierie des Données, spécialisée en Data Science, IA et Machine Learning. Passionnée par l'analyse et la modélisation de données, je maîtrise tout le cycle de vie des données, de l'exploration à la modélisation prédictive, en passant par le nettoyage et la visualisation. Forte de compétences en NLP et en développement d'algorithmes de machine learning, je recherche un stage de fin d'études à partir de février 2025 pour mettre en pratique mes compétences dans des projets d'innovation

## Education

**École Nationale des Sciences Appliquées, Khouribga**, Informatique et Ingénierie des Données. Sept 2020 – May 2025  
- Cycle d'ingénieur en informatique et ingénierie des données, 2022 – 2025  
- Années préparatoires intégrées, 2020 – 2022  
**Lycée Ibn Abdoun, Khouribga**, Baccalauréat Science physique. Sept 2019 – May 2020

## Expériences Professionnelles

**Data Science Intern | Estimation des niveaux d'obésité**, HubbleMind Octobre 2024 – Novembre 2024

- Réalisation d'un projet intitulé "Estimation of Obesity Levels Based on Eating Habits and Physical Condition", en utilisant plusieurs techniques de machine learning.
- Exploratory Data Analysis (EDA), prétraitement des données et mise en place d'un pipeline complet de modélisation.
- Développement et entraînement de quatre modèles de machine learning : Logistic Regression, Random Forest, Gradient Boosting, et MLP Classifier.
- Le modèle Gradient Boosting a obtenu la meilleure précision avec 96.89%, tandis que le MLP Classifier et Random Forest ont atteint des précisions supérieures à 94%.
- Logistic Regression a également donné des résultats satisfaisants avec une précision de 73.68%.

**Stage PFA | Maintenance prédictive des équipements industriels | Projet d'automatisation de la détection des défauts de roulements**, Groupe OCP Juillet 2024 – Septembre 2024

- Analyse et prétraitement de données vibratoires provenant de capteurs industriels.
- Développement d'un modèle de clustering (k-means) pour la classification des vibrations.
- Mise en place d'un modèle d'Isolation Forest pour détecter les anomalies.
- Création d'une interface graphique avec Tkinter pour automatiser tout le processus.

**Stage d'initiation | Développement d'une application de gestion des stagiaires**, Groupe OCP Juillet 2023 – Août 2023

- Développement d'une application Java avec Swing pour la gestion des stagiaires.
- Conception d'une interface graphique intuitive pour la gestion des stagiaires.
- Optimisation de la base de données pour des requêtes plus rapides et sécurisées.

## Certifications

---

- Advanced Learning Algorithms – Coursera
- Supervised Machine Learning: Regression and Classification - Coursera
- Agile Project Management – Coursera
- Python – OpenClassroom

## Projets académiques

---

### Chatbot de recherche d'emploi et de stage

- Développement d'un chatbot utilisant le machine learning pour aider à la recherche d'offres d'emploi et de stage. Ce chatbot comprend les requêtes en langage naturel et propose des résultats personnalisés.
- Amélioration de l'efficacité des recherches de 70% grâce à une meilleure compréhension des besoins utilisateurs.
- Technologies: Rasa, React, Node js, Machine Learning, Python

### Détection de Fraude avec des Modèles de Machine Learning

- Projet de détection de fraudes en utilisant des algorithmes de machine learning sur un jeu de données fortement déséquilibré. Plusieurs modèles, dont l'Isolation Forest, le Classifieur SGD (gradient descent stochastique) et un Réseau de Neurones, ont été entraînés et comparés pour identifier les transactions frauduleuses.
- Précision du modèle Neural Network atteignant 92% sur des transactions frauduleuses.
- Technologies: Python, Scikit-learn, Keras, Jupyter Notebook, Pandas, Matplotlib

### Analyse de données pour la détection du diabète

- Technologies: Talend, Visual Studio, SSMS, Power BI

### Nettoyage et analyse de données de ventes pour la prise de décision

- Technologies: Python, Power BI

### Application de location de voitures

- Technologies: Java - JEE avancée - MySQL

### Application de digitalisation des fonctionnalités d'arrondissement

- Technologies: Java - JEE avancée – MySQL- Spring

## Bénévolats

---

**Membre actif de la cellule Exposition lors de l'événement DataVerse**

**Membre actif du club MI (Mathématique et Informatique)**

**Participante à la compétition Hult Prize**

## Technologies

---

**Langages de programmation :** Python,R, Scala, C, Java

**Big Data & Machine Learning :** Hadoop, Spark, Scikit-learn, Keras, TensorFlow, Pytorch, NLP, Traitement d'image

**Développement Web :** HTML, CSS, JavaScript, TypeScript

**Bases de Données :** MySQL, NoSQL

**Outils & Environnements :** Linux/Ubuntu, Git, GitHub, Power BI, Excel, Logiciels Bureautiques (Word, PowerPoint, etc.)

## Langues

---

**Arabe:** Langue maternelle

**Français:** Courant, lu, écrit et parlé

**Anglais:** Intermédiaire, lu, écrit et parlé