

# YOSRA KAROUI


Ingénieure Informatique | DATA Scientist

✉ yosra.karoui@esprit.tn |  LinkedIn |  Github |  Website


## COMPÉTENCES

- **Langages de programmation** : Java , Python , Java Script , Scala , R , SQL
- **Bases de données** : MySQL , PostgreSQL , MongoDB
- **Big Data** : Cloudera , Hadoop , Kafka , Spark
- **Bibliothèques Machine Learning et IA** : PyTorch , XGBOOST , TensorFlow , Scikit Learn
- **NLP** : NLTK , SpaCy , Transformers
- **Analyse de données** : PowerBI , Tableau , Data Perspective Studio
- **Manipulation de données** : Pandas , Numpy , Matplotlib , Seaborn
- **Technologies Cloud** : Microsoft Azure , AWS
- **DevOps** : Jenkins , Docker , SonarQube , Grafana , Prometheus , Git , MLflow
- **ETL** : Apache Airflow , SSIS , Talend
- **Data Warehousing** : Snowflake , BigQuery
- **Linux Et Shell Scripting**
- **Languages** : E nglish(C1) , French (C1 ) , Spanish (A1)

## EXPÉRIENCE

- **Dassault Systèmes** 

05/ 2024 - 02/08  
R&D Data Engineer - Stage de fin d'études  
Vélizy-Villacoublay, France

  - Conception d'un pipeline de données en temps réel pour l'intégration de bases de données locales dans la plateforme 3DEXPERIENCE.
  - Création de tableaux de bord dans Data Perspective Studio pour fournir des insights en temps réel.
  - Développement d'un prototype de chatbot avec fine-tuning de Transformers pré-entraînés .
- **ESPRIT-Tech** 

06/2023 - 09/2024  
Data Scientist - Stage  
Tunis - Tunisie

  - Implémentation de réseaux de neurones convolutifs pour un système de débruitage d'images .

## EDUCATION

- **École supérieure privée d'ingénierie et de technologie (Bac+5)**


09/2024  
Diplôme d'ingénieur en informatique avec specialisation en Data Science  
Tunis, Tunisie

## PROJETS ACADÉMIQUES CLÉS

- **Projet DevOps**


11/2023 - 12/2023

Outils: [Jenkins / Nexus / Docker / Git / Springboot / SonarQube / Prometheus / Grafana]


  - Rationalisation du pipeline CI/CD pour un backend Spring Boot avec automatisation du code, contrôle qualité, déploiement Docker, surveillance et alertes par e-mail.
- **MindMatch : Career Center Platform**


02/2023-05/2023

Outils: [BeautifulSoup / SpaCy / NLP / KMean / Logistic Regression ]


  - Développement d'une plateforme de centre de carrière avec extraction de données (web scraping), gestion des données (Python) et algorithmes de machine learning pour profilage, régression et classification. Création d'un tableau de bord Power BI pour visualiser les tendances du marché du travail.
- **Prédiction de la maladie rénale chronique**

10/2022-12/2022

Outils: KNN / KMean / SVM / Naive Bayes


  - nettoyage, visualisation et préparation des données.
  - Comparaison de modèles de classification : KNN, Naive Bayes, SVM, Random Forest, Arbre de décision..