

Data Scientist / Engineer

Ingénieur en informatique passionné par l'IA et la data.
Spécialisé dans l'analyse et le traitement de données
ainsi que le développement de systèmes d'IA et de Data.

Expériences professionnelles

- Depuis février 2024
- **Développeur Data Indépendant** Paris
 - Développement de scripts d'extraction de données et constitution de bases de données analytiques.
 - Réalisation de traitements et d'analyses approfondies sur des corpus de textes.
 - Amélioration de la qualité de réponse d'un LLM de 30% grâce à l'implémentation d'une architecture RAG.
 - **Outils** : Python, Tensorflow, Matplotlib, Haystack, MongoDB, Playwright.
 - De novembre 2022 à septembre 2023
 - **Consultant Data Scientist Sogeti part of Capgemini** Issy-les-Moulineaux (en alternance)
 - Conception et mise en œuvre de modèles de NLP descriptifs et prédictifs et de pipelines Azure Machine Learning.
 - Création de tableaux de bord analytiques et d'applications.
 - Accompagnement des clients dans la valorisation de leurs données, générant une augmentation de 20% des conversions.
 - **Outils** : Python, Scikit-Learn, HuggingFace, Tensorflow, Azure Machine Learning, Streamlit, Docker.
 - De novembre 2021 à août 2022
 - **Stagiaire en Deep Learning ESI** Alger
 - Développement d'une architecture Deep Learning pour la reconnaissance optique de numéros plaques d'immatriculation.
 - Fine-tuning des modèles et Déploiement de l'architecture sur une machine virtuelle GCP.
 - Développement d'une API Rest avec un endpoint pour de l'inférence en temps réel.
 - **Outils** : Python, openCV, Pytorch, GCP, Flask, Weights & Biases.

Diplômes

- D'octobre 2021 à septembre 2023
- **Master en Informatique, Parcours Machine Learning for Data Science (MLSD) Université Paris Cité** Paris
 - Programme: Big Data Mining, NLP, Deep Learning, Times Series, Recommender Systems.
 - De septembre 2017 à septembre 2022
 - **Diplôme d'ingénieur en informatique École nationale Supérieure d'Informatique ESI ex INI** Alger
 - Programme: Calcul Distribué, SGBD avancés, Structures de données et de fichiers, Analyse de données, Recherche Opérationnelle.

Projets Personnels

TweetAnalysis : plateforme web d'analyse de tweets en temps réel (*trends, sentiments, visualisations, etc.*)

- Python, Pandas, Streamlit, ScikitLearn, nltk, NLP, Machine Learning.

Compétition Kaggle : architecture deep learning pour la prédiction des périodes de sommeil chez l'enfant.

- Python, Tensorflow, Time Series, LSTM, CNN, BiLSTM, Matplotlib, Pandas, Parquet.

Pipeline ELT: Création de tables de faits et de dimensions à partir de tables sources.

- SQL, dbt, Snowflake, Airflow.

Compétences

Data Stack

Airflow, Snowflake, BigQuery, Spark, dbt, Java

Data Analytics

SQL, Looker, Streamlit, NoSQL DBMS (MongoDB)

Data Science & Machine Learning

Python, R, Pandas, Numpy, Matplotlib, Tensorflow, Sk-Learn, MLFlow, MLOps, LLM & RAG

Infrastructure & Software

Docker, Bash, git, REST APIs (*Flask, FastAPI*), Design Patterns

Cloud & DataOps

- Certifié **Microsoft Azure Data Scientist Associate**
- Certifié **Databricks Lakehouse Fundamentals**
- En cours de certification **Google Cloud Professional Data Engineer**.

Centres d'intérêt

Jeu d'Échecs, Cinéma, Voyage

Langues

Français (C2), Anglais (C2)

Réseaux sociaux

in @lounes-mechouek

🐱 @lounesmechouek