

SAAD YAQINE

Data Scientist

@ saadyaqine91@gmail.com

📞 +33 7 66 16 96 16

in Saad Yaqine

✉️ Saad Yaqine

FORMATION

Cycle ingénieur en génie informatique

École Polytechnique de Marseille

EXPÉRIENCES PROFESSIONNELLES

Data Scientist

Bouygues Telecom : Mars 2024 - Mars 2025 – Meudon, France

- Développement d'un modèle de classification réduisant de 46% les appels inefficaces avec 94% de précision.
- Automatisation du déploiement via pipelines MLOps (Docker, GCP), réduisant le time-to-market de 30%.
- **Stack :** Python, PySpark, Dataiku, Vertex AI, BigQuery, Teradata

Machine Learning Engineer

LebonCoin: Avril 2023 - Octobre 2023 – Paris, France

- Classification automatique d'annonces avec TF-IDF et embeddings, atteignant 92% de précision.
- Déploiement du pipeline NLP via microservice FastAPI pour automatisation sur GCP.
- **Stack :** Python, Vertex AI, FastAPI, PostgreSQL, MLFlow

Data Scientist

Université polytechnique de Valence : Décembre 2022 - Mars 2023 – Valence, Espagne

- Construction de datasets multilingues avec modèles LLM (AraGPT, BLOOM) pour génération de texte.
- Développement d'un modèle de détection automatique de langue avec 90% de précision.
- **Stack :** Python, PyTorch, Transformers, Scikit-learn, Git

Data Engineer

Marsa Maroc: Juillet 2022 - Aout 2022 – Mohammadia, Maroc

- Analyse exploratoire des données opérationnelles pour maintenance prédictive.
- Pipeline ETL et visualisations Tableau, réduisant de 40% le temps du reporting mensuel.
- **Stack :** Python, SQL, Snowflake, Tableau

PROJETS

FinSentBot - Système de Trading Automatique

- Architecture temps réel avec Kafka collectant 18K+ prix d'actions pour analyser sentiment financier et générer signaux de trading
- Modèle PyTorch de classification avec précision de 90%
- **Stack :** Python, Kafka, PyTorch, FinBERT, yfinance, Streamlit, BeautifulSoup, Docker



À PROPOS DE MOI

Ingénieur en Data Science spécialisé dans le développement de modèles de Machine Learning et l'analyse de données à grande échelle. Passionné par la résolution de problèmes complexes et l'optimisation des processus métiers grâce à des solutions data innovantes dans divers secteurs (télécommunications, e-commerce, industrie).

COMPÉTENCES

- **Langages & Données :** Python, R, SQL, C++, SAS, PySpark, Pandas, Scikit-learn, Dataiku, Airflow, Snowflake, Databricks.
- **Cloud & IA :** AWS, Azure, Docker, Kubernetes, GitHub, Power BI, Streamlit, FastAPI, PyTorch, TensorFlow, LangChain, Kafka.
- **Machine Learning :** Scikit-learn, Transformers, réseaux de neurones, apprentissage supervisé, fine-tuning.
- **Visualisation & BI :** Tableau, Power BI, Streamlit, analyse de signaux, traitement en temps réel.

CERTIFICATIONS

- **MLOps Practitioner & Advanced Designer** (Dataiku)
- **Google Cloud Professional Data Scientist** (En cours)

LANGUES

- **Français :** Courant
- **Anglais :** Courant (TOEIC 900/990)

CENTRES D'INTÉRÊT

- Machine Learning et Deep Learning appliqués à la résolution de problématiques métiers multisectorielles
- Architecture de données temps réel et systèmes de streaming (Kafka, Apache Spark)
- NLP financier et modèles de traitement du langage naturel (FinBERT, Transformers)
- MLOps et déploiement automatisé de modèles en production