



Zakaria TOZY

Data/ML Engineer | CDI

☎ 0617407077 | ✉ zakaria.tozy@icloud.com | in zakaria-tozy | 📍 Paris | 🌐 github.com/zack242

Ingénieur en Data Science avec une double formation (École Polytechnique/ECE Paris), spécialisé en Machine Learning et Data Engineering. Je recherche un poste en CDI pour contribuer à des projets innovants.

COMPETENCES

Machine Learning : TensorFlow, Scikit-Learn, Deep Learning, NLP

Cloud : GCP, AWS, Azure, Docker

Programmation : Python, SQL, Java

Soft Skills : Autonomie, Rigueur, Esprit d'analyse, Apprentissage rapide

Langues : Français (Bilingue), Arabe (Bilingue), Anglais (Courant - TOEIC 875)

Data & Cloud Engineering : PySpark, Databricks, ETL, Data Pipelines, dbt

FORMATIONS

Institut Polytechnique de Paris (École Polytechnique)

Janvier 2024

Master 2 en Data Sciences

Mention Très bien

- Spécialisation : Deep Learning, NLP, Machine Learning, Cloud Computing

École Centrale d'Électronique (ECE Paris)

Janvier 2024

Ingénieur en Systèmes d'Information et Big Data Analytics

Top 5% de la promotion

- Spécialisation : Data Engineering, Machine Learning, Statistiques Avancées

EXPERIENCES

AXA Investment Managers - AXA IM

Août 2023 — Janvier 2024

Data Engineer/Data Analyst Intern

Paris

- Conception et mise en place de pipelines de données sur Databricks pour l'ingestion et la gestion des données métiers
- Optimisation du modèle de données et migration des pipelines de distribution
- Tests de nouveaux clusters Databricks et création d'un PoC pour l'anonymisation des données
- Création de tableaux de bord et calcul des KPIs financiers

Kalima Blockchain et IoT

Avril 2022 — Août 2022

Software Engineer Intern

Paris

- Développement d'un explorateur de blockchain propriétaire en Java/LevelDB (1000+ TPS)
- Implémentation d'algorithmes d'analyse de données pour le monitoring des transactions
- Conception d'un système d'authentification sécurisé

Le Crédit Lyonnais - LCL

Janvier 2020 — Février 2020

IT Intern

Paris

- Étude de migration Microsoft Office 2010 vers 2016 et analyse de son impact sur les processus métiers
- Développement et exécution de procédures de tests sur un parc de 23 000 postes de travail

PROJETS

Data Challenge Natixis

- Analyse des données textuelles de la FED avec FinBERT dans le but de prédire son impact sur le marché

Bitcoin Analysis

- Pipeline de données (Python, SQL, Airflow, dbt, Snowflake) pour l'analyse de 1M+ transactions Bitcoin par jour

Twitter Sentiment Analysis

- Pipeline temps réel utilisant Kafka et un KNN pour la classification des sentiments en temps réel et en topic

Désagrégation de la Courbe Énergétique Capgemini

- Extraction de la composante thermosensible de la consommation énergétique avec un modèle statistique et un réseau de neurones