



Zakaria TOZY

Data Engineer

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

Ingénieur en Systèmes d'Information et Data Science avec une double formation (ECE Paris/Polytechnique), passionné par les données et les nouvelles technologies. Je recherche un poste en CDI pour contribuer à des projets innovants et à forte valeur ajoutée.

COMPETENCES

Programmation : Python, SQL, Java, C++, C#

Machine Learning : Scikit-Learn, NLP, Computer Vision, Deep Learning

Devops : Docker, CI/CD, Git, Scrum

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

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

Data & Cloud Engineering : PySpark, BigQuery, ETL, Data Pipelines, Databricks, Snowflake, GCP, dbt

FORMATIONS

École Centrale d'Électronique (ECE Paris)

Janv 2024

Ingénieur en Systèmes d'Information et Big Data Analytics

Top 5% de la promotion

- Spécialisation : Programmation, Bases de données, Électronique, DevOps

Institut Polytechnique de Paris (Ecole polytechnique)

Janv 2024

Master 2 en Data Sciences

Mention Très bien

- Spécialisation : Machine Learning, Big Data, Cloud Infrastructure, Deep Learning, NLP

EXPERIENCES

AXA Investment Managers - AXA IM

Août 2023 — Janv 2024

Data Engineer/Data Analyst Intern

Paris

- Conception et mise en place d'une dizaine de pipelines de données pour l'ingestion de données métiers.
- Optimisation du modèle de données et migration des pipelines de distribution utilisés pour le reporting financier de 844 milliards d'euros d'actifs en 2023.
- Tests de nouveaux clusters sur Databricks et création d'un PoC (Proof of Concept) pour l'anonymisation des données sensibles.
- Création de tableaux de bord pour les contrôleurs financiers.
- Calcul des indicateurs de performance KPIs pour la vision distribuée (ME, OV, CIS, FX).
- Travail sur la qualité et la validation des données pour le reporting.

Kalima Blockchain et IoT

Avr 2022 — Août 2022

Software Engineer Intern

Paris

- Développement d'un explorateur de blockchain en Java/LevelDB (+1000 transactions par seconde).
- Implémentation de nouvelles fonctionnalités d'administration pour la blockchain Kalima en Node.js.
- Conception et développement d'un système d'authentification multisignature pour les DApps (Decentralized Applications).

Le Crédit Lyonnais - LCL

Janv 2020 — Fév 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

Bitcoin Analysis

- Création d'un pipeline de données avec Python, SQL, Airflow, dbt et Snowflake pour l'analyse des transactions Bitcoin, traitant plus de 1 million de transactions par jour.

Twitter Real-time Analysis

- Développement d'un système de visualisation en temps réel des flux Twitter par topics à l'aide de Kafka et de l'algorithme KNN (K-Nearest Neighbors).

Energy Disaggregation Capgemini

- Extraction de la composante thermosensible de la consommation énergétique pour prévoir la demande future en France.