

ELGHAZI SOUFIANE

Junior Data Engineer

@ soufiane.el-ghazi@outlook.com soufiane-el-ghazi +212 6533-61122 SoufianeElghazi
Casablanca, Maroc elghazisoufiane.vercel.app/ M elghazisoufiane02 </> Soufiane-EL



PROFIL

”Les données racontent une histoire, encore faut-il savoir l’écouter.”

Data Engineer avec plus d’un an d’expérience, je suis spécialisé dans la transformation de données brutes en informations structurées et exploitables. Maîtrisant **Python**, **SQL** et **Spark**, je conçois et optimise des **pipelines ETL**, développe des **architectures de données** haute performance et relève des défis complexes avec une approche analytique précise. Toujours en quête d’innovation, je m’adapte rapidement aux **technologies émergentes** pour fournir des solutions de données fiables.

Alliant expertise en ingénierie et analytique avancée, je mets en œuvre des **modèles d’apprentissage automatique** tout en maintenant une infrastructure de données robuste. Mon expérience s’étend du développement de **pipelines d’analyse en temps réel** à la création de **tableaux de bord interactifs**, garantissant des solutions évolutives.

FORMATION

Ingénierie des Données et des Connaissances

École des Sciences de l’Information [ESI]

Octobre 2021 – Juin 2024

Rabat

Classes Préparatoires

Lycée Ibn Tahir

Septembre 2019 – Juin 2021

Errachidia

Baccalauréat en Mathématiques B

Lycée Technique Ibn Al Haitam

Septembre 2018 – June 2019

Ouarzazate

EXPÉRIENCE PROFESSIONNELLE

Consultant Data

ENOVA RT • Temps plein • Sur site

Septembre 2024 – Présent • 11 Mois

Rabat

- Implémentation d’un entrepôt de données centralisé pour les opérations BI avec processus ETL, modélisation dimensionnelle et documentation complète
- Suggestion et développement d’une application de matching de CV utilisant le NLP pour améliorer le processus de recrutement
- Proposition et implémentation d’un chatbot local pour la documentation utilisant RAG et Ollama pour optimiser le partage des connaissances
- Développement de scripts SQL et PL/SQL pour l’injection, l’extraction et la mise à jour des données
- Conception et déploiement de tableaux de bord dynamiques pour les métriques hospitalières, de facturation et de paiements

Python PowerBI PostgreSQL SQL Server Talend RAG Java Fastapi

Ingénieur Vision par Ordinateur & IA

FIVERR • Freelance • À distance

Mai 2024 – Septembre 2024 • 4 Mois

Paris

Détection et Extraction de Cartes

- Développement d’une fonctionnalité de vérification pour une application mobile utilisant Flutter, OCR pour l’extraction de texte, et OpenCV pour l’amélioration d’image, focalisée sur la validation de la carte d’identité nationale (CIN)
- Collaboration avec le client pour définir les exigences du projet et fournir des solutions de validation adaptées

Python OpenCV Flask React OCR Yolo

Stagiaire Data Science PFE

 ENOVA RT • Stage • Sur site

 Février 2024 – Août 2024 • 6 Mois

 Rabat

Modélisation Prédictive des Infections Nosocomiales

- Développement d'un modèle prédictif des infections nosocomiales avec une précision de 92%, déploiement dans une application web et automatisation des processus ETL/ML
- Implémentation d'un pipeline de prédiction en temps réel, réduisant le temps de réponse aux infections de 30% et conception d'un modèle de classification de 50 types de cancer avec une précision de 88%

Python Kafka Spark Airflow Django Celery Redis

PROJETS

Outil d'Extraction des États Financiers :



Développement d'un outil automatisé pour extraire et afficher les états financiers à partir de rapports financiers en PDF

- Défi :** Extraire des données financières clés à partir de PDF complexes

OCR OpenCV PyTesseract Flask React

Prédiction du désabonnement des clients en Temps Réel:



Traitement de données en temps réel pour l'analyse prédictive de l'attrition client

- Défi:** Gestion des flux de données à haute vitesse tout en maintenant la précision des prédictions

Spark Kafka Power BI Snowflake

COMPÉTENCES

Langages de Programmation

Python SQL Java DAX

Développement Backend

FastAPI Flask

Bases de Données

PostgreSQL Elasticsearch MongoDB Snowflake Cassandra SQL Server

Big Data & Cloud Computing

Spark PySpark Hadoop Kafka Azure

ETL & Data Modeling

Talend DBT Dimensional Modeling Airflow

DevOps & CI/CD

Docker Linux Git

Visualisation & BI

Power BI Dashboards

Machine Learning / Deep Learning


Computer Vision NLP LLMs RAG Ollama

Compétences Personnelles

Résolution de problèmes Communication Adaptabilité Gestion du temps

CERTIFICATIONS

 **Microsoft Azure Fundamentals (AZ-900)**
 **Microsoft Power BI Data Analyst Associate (PL-300)**

 **Microsoft Azure Data Fundamentals (DP-900)**
 **Microsoft Fabric Data Engineer Associate (DP-700)**

LANGUES

- **Arabe:** Langue maternelle
- **Anglais:** Courant (Parlé, Lu, Écrit)
- **Français:** Courant (Parlé, Lu, Écrit)

MES RÉALISATIONS

 **My Microsoft Power BI Udemy Course**
 **Understanding Power BI's Engines**
 **Data Engineering Concepts Article**

 **My Learn Git Udemy Course**
 **Understanding DAX Contexts**