



Mohammad ZENNOU

Data Engineer | GCP Certified - Professional Data Engineer



✉ zennou1md@gmail.com

☎ +33 6 65 75 47 35

🏠 Lille

📍 France

## À propos de moi

Cloud Data Engineer avec une expertise dans la transformation de données brutes en insights précieux. Compétences en Python/Scala, SQL et cloud (GCP, AWS). Esprit d'équipe et de collaboration, capable de fournir des solutions de données évolutives pour répondre aux besoins de l'entreprise.

## Expériences professionnelles

### Data Engineer

Exomind/Decathlon

05-2023 – 10-2023

Lille

Conception et création de flux de données pour l'équipe de développeurs backend du projet ETCO de Decathlon, facilitant le traitement efficace des données et l'amélioration de la performance du système.

- Utilisation avancée de **Databricks** pour explorer et analyser les données.
- Conception et mise en œuvre de nouveaux pipelines de données à l'aide de **Spark** et **Scala**, permettant une gestion plus efficace des données et une augmentation significative de la vitesse de traitement.
- Migration réussie du système de gestion de tâches **Jenkins** vers **Apache Airflow**, améliorant la planification des tâches, la surveillance et la gestion des flux.

### Cloud Data Engineer

JEMS Group

01-2023 – 04-2023

France

Mise en œuvre d'une solution de bout en bout pour la collecte, la transformation et l'analyse de données en batch et en temps réel provenant de capteurs automobiles, à la fois on-premise et dans le Cloud à l'aide de Google Cloud Platform (**GCP**).

- Ingestion de données via le service **Pub/Sub**
- Pipeline de données et traitement des données sous **DataFlow** avec **Apache Beam** (SDK Python)
- Stockage des données brutes sur l'entrepôt de données **BigQuery**
- Construction de modèles de données cohérents et standardisés avec l'outil **Dbt**  
Création de Dashboard et de rapports en temps réel avec **Looker Studio**
- Gestion de code et orchestration avec de **Terraform**, **Github Action**

### Data Scientist

Bayer AG

04-2022 – 10-2022

Allemagne

Concevoir et mettre en œuvre un modèle NLP pour aider les experts en réglementation à trouver des informations exploitables parmi des milliers de documents provenant des agences sanitaires

- Collectant et étiquetant les données avec **AWS Ground Truth**, et **AWS S3**.
- Transformation de la données avec **AWS Glue**.
- Nettoyage et analyse de l'ensemble des données avec **Python**, **Pandas**, **Numpy**, **Matplotlib**, **Seaborn** avec **AWS SageMaker**.
- Création d'un modèle pour la modélisation thématique en utilisant **BERTopic**, **Transformer**, **Flair**.
- Création d'un modèle pour la classification en utilisant **Pytorch**, **Sklearn**.
- Développement d'une application web avec **Stramlit**.

Résultats: 90% de précision.

### Data scientist

Henceforth

03-2020 – 10-2020

Maroc

Conception et mise en œuvre d'un modèle de reconnaissance vocale SR permettant d'identifier une personne à partir de sa propre voix.

- Collecte des fichiers audios et constitution de l'ensemble de données en utilisant des scripts en **python**.
- Prétraitement du jeu de données (Voice Activity Detection, Reduire le bruit, enlever le silence, etc...) en utilisant **pyaudio** et **Librosa** et des fonction en python.
- Extraire les caractéristiques de chaque fichier audio (MFCC, LPC) en utilisant **Librosa**.
- Entraînement de plusieurs modèles d'apprentissage automatique (SVM, RandomForest, ANN, et CNNs) pour obtenir le meilleur classificateur, implémentation des modèles avec **Sklearn** et **Tensorflow**.
- Développer une application Web avec **Streamlit** et **Reactjs** pour la déployer et la mettre en production.

Résultats: 98% de précision.

Compétences Techniques

Langages de programmation

Python, Scala, Java, Bash, SQL

Big Data

Spark, Hadoop, Databricks, HDFS/MapReduce, Dbt

Data Science

NumPy, Pandas, Matplotlib, , Scikit-learn, OpenCV, Librosa, NLTK, Spacy, Transformer

Business Intelligence

Power BI, Tableau, Locker, Data Studio

Base données

MySQL, MongoDB, ElasticSearch, BigQuery, BigTable.

Frameworks

Fast API, Sklearn, Tensorflow, Pytorch, Streamlit, Spring Boot, Angular 11, Bootstrap

Outils

Git/Gitlab, Terraform, Docker, Kubernetes, Airflow, GCP, AWS, Jira

Certifications

●	<b>GCP Professional Data Engineer</b> Google Cloud <i>05-2023</i>	<i>GCP</i>
●	<b>Machine Learning Engineering for Production (MLOps) Specialization</b> Google & Deepai <i>04-2023</i>	<i>Coursera</i>
●	<b>DevOps on AWS Specialization</b> AWS <i>04-2023</i>	<i>Deepai</i>

Diplômes et Formations

●	<b>BootCamp Cloud Data Engineering</b> DataScientest/ MinesParisTec <i>03-2023 – 06-2023</i>	<i>Paris</i>
●	<b>M2 Intelligence Artificielle</b> Université d'Artois <i>08-2021 – 08-2022</i>	<i>France</i>
●	<b>M1/M2 Business Intelligence et Big Data Analytics</b> UCD <i>08-2018 – 08-2020</i>	<i>Maroc</i>
●	<b>L3 Genie Logiciel</b> FSTE <i>10-2017 – 07-2018</i>	<i>Maroc</i>

Centres d'intérêt

Voyage, AI, Cuisine