



# Wacim BRIANE

+33 695 75 40 45 — wacim.br@gmail.com — <https://www.linkedin.com/in/wacim-briane-379916229/> — France métropolitaine

**Data Engineer** — À la recherche d'un nouveau défi en data engineering

## Expérience

### Orange France

Mars 2024 – Août 2024

#### Data Scientist

- Collaboration avec les équipes métier pour formaliser un besoin en tâche de classification.
- Création d'une pipeline de traitement automatisée via BigQuery, nettoyage et transformation des données.
- Réalisation d'une analyse exploratoire (EDA) pour détecter les tendances et les anomalies.
- Prétraitement avancé : vectorisation, gestion des classes rares, et tokenisation des textes.
- Labellisation manuelle et semi-automatique des données pour enrichir les jeux d'entraînement.
- Entraînement et comparaison de modèles BERT, CamemBERT et DeBERTa avec Hugging Face.
- Fine-tuning des modèles pour optimiser les scores F1, précision et rappel.
- Évaluation des performances avec des métriques personnalisées et courbes ROC.
- Analyse des biais et risques éthiques, rédaction d'un rapport de conformité.
- Déploiement du modèle sur GCP avec scripts de monitoring pour la production.

### Carrefour Administrative France

Mars 2023 – Août 2023

#### Data Engineer - Analyst

- Compréhension des besoins métier et traduction des règles de gestion pour les livrables.
- Analyse des données SAP, création de vues et transformation pour faciliter l'interprétation.
- Extraction de données via SQL et PySpark depuis différentes sources (Google Drive, BigQuery).
- Création et automatisation de pipelines ETL avec cron, Bash/Python pour assurer la fiabilité.
- Optimisation des performances des requêtes et des transformations.
- Collaboration avec les équipes Data pour garantir la qualité des données avec contrôles et tests unitaires.
- Création de dashboards interactifs dans Looker Studio pour le suivi des KPIs.

## Formations

### Master 1 et Master 2 en Data Science

INSA Rouen - Université de Rouen

2022 – 2024

Rouen, France

### Licence en Informatique, parcours Data Science

Université de Rouen

2021 – 2022

Rouen, France

### Master 1 en Business Intelligence

Université des Sciences et de la Technologie H.B

2020 – 2021

Alger, Algérie

## Projets

### - Traitement de fichiers bruts et création de tables optimisées sur Databricks.

(PySpark, Spark SQL, Databricks, Delta Lake)

Nettoyage de données hétérogènes, création de colonnes normalisées, agrégation métier, et export final en fichiers Parquet.

### - Conception d'un pipeline ETL orienté objet pour le traitement de données e-commerce.

(Python, POO, Bash, Logging)

Développement de plusieurs classes pour chaque étape du pipeline, implémentation de tests unitaires, gestion des erreurs et création de logs pour le suivi des exécutions.

## Certifications

- Google Cloud Fundamentals: Core Infrastructure.
- Build your first Machine Learning Pipeline using Dataiku.
- PySpark - Apache Spark Programming in Python.

## Compétences

- SQL (PostgreSQL), Spark, Bash, Scala, Git, Excel, Power BI, Databricks, Dataiku, CI/CD, Hadoop, Agile, ETL, Terraform, Snowflake, Kafka.
- Python (Scikit-learn, Numpy, Pandas, Pytorch).
- Machine learning, Deep learning, NLP, LLM.
- Google Cloud Platform (BigQuery, Sheets, Google Colab, Google Data Studio, AppScript).

### Soft skills :

Autonomie, Adaptation, Communication, Rigueur,

### Centres d'intérêt :

Musculation, Football, Cuisine.

### Langues :

Français , Anglais.