



LINA BENZEMMA

DATA SCIENTIST/DATA ENGINEER

✉ lina.benzemma@yahoo.com

☎ +33 7 49 59 85 14

📍 Île-De-France



[Github](#)



[Portfolio](#)



[Linkedin](#)

Actuellement en Master 2 MOSEF Data Science & Big Data à l'Université Panthéon-Sorbonne, en alternance chez Sanofi, je suis à la recherche d'un CDI à partir de septembre/octobre 2025 dans le domaine de la data science ou du Big Data. Je souhaite intégrer une structure dynamique où je pourrai mobiliser mes compétences en analyse de données, machine learning, ingénierie des données et développement d'applications analytiques.

Expériences Professionnelles

SANOFI-DATA SCIENTIST (ALTERNANCE)

2024-2025

Vitry-sur-Seine, France

- En charge du développement et de l'amélioration de deux applications R Shiny.
- Développement de logiciels
- Gestion de projet
- Analyse statistique et méthodes statistiques

IFP- INGÉNIEURE MACHINE LEARNING (STAGE)

2024 (5 MOIS)

Rueil-Malmaison, France

- Développement de méthodologies de machine learning sur des données de glissements de terrain.

AUCHAN -GESTION DU COFFRE CENTRAL (JOB ÉTUDIANT)

2021-2024

Le Kremlin-Bicêtre, France

Formation

Master 2 – Modélisations Statistiques Économiques et Financières (MOSEF)

2024-2025

Université Panthéon Sorbonne

Paris, France

- Compétence dans les environnements big data (Microsoft Azure, Hadoop, Spark) et le web scraping.
- Expertise en outils de programmation (Python, Scala, Java, Linux)
- Maîtrise des méthodes d'apprentissage statistique (Machine Learning, Deep Learning)

Master 1– Mathématiques Appliquées et Apprentissage Statistique

2023-2024

Université Paris Saclay

Versailles, France

- Compétences en optimisation numérique, inférence statistique, modélisation linéaire, et machine learning.
- Expertise en algorithmique randomisée, recherche opérationnelle, et traitement de données.

Licence – Mathématiques Appliquées

2020-2023

Sorbonne Université

Paris, France

Projets académiques

TrioVision – Classification d'images géolocalisées avec EfficientNet:

Modèle EfficientNet-Bo entraîné sur le Google Landmarks Dataset v2 pour prédire le pays d'origine d'une image.

Prédiction de la Volatilité des Cryptomonnaies :

Prédiction de la volatilité des cryptomonnaies et optimisation de portefeuille via des modèles GRU, LSTM, TCNN et une API FastAPI.

Real Time Data Streaming with Apache Kafka on GCP :

Développement d'une plateforme de streaming avec Kafka, NiFi, GCS, BigQuery ML, Elasticsearch et Kibana pour le suivi de trajets de taxis, l'estimation dynamique des prix et le clustering géographique en temps réel

Template Scala Spark – Traitement de données modulaire :

Lecture, traitement et écriture de données (CSV, Parquet, Hive) via un pipeline modulaire en Scala et Apache Spark.



HARD SKILLS :

Python, Rshiny, Scala, SQL, HTML, Apache Spark, Nifi, Kafka, GCP, Hadoop, Hive, Docker, Tableau, Dataiku.



SOFT SKILLS:

Esprit analytique, adaptabilité, curiosité, gestion du temps et organisation



LANGUES:

Français, Anglais (C1)



CENTRES D'INTÉRÊT :

Randonnée, Voyages, Lecture, Cinéma et la peinture