



MOUNA OUATTARA

DATA SCIENTIST

CONTACT

✉ mounaouattara04@gmail.com

☎ +330758286090

📍 Palaiseau, Île-de-France, France

🌐 Portfolio

🌐 LinkedIn

🌐 GitHub

COMPÉTENCES

LANGAGES & DÉVELOPPEMENT

Python R SQL TypeScript

Java

Bash.

DATA SCIENCE & ML

Pandas

NumPy

Scikit-learn

XGBoost

TensorFlow

Keras

PyTorch

MLOPS & INFRASTRUCTURE

Docker

MLflow

Airflow

AWS (S3)

AWS(SageMaker)

Databricks

BASES DE DONNÉES

PostgreSQL

MySQL

MongoDB

VISUALISATION & BI

Power BI

Matplotlib

Seaborn

Plotly

Streamlit.

CERTIFICATIONS

🏆 AWS Certified Machine Learning Specialty (en cours)

LANGUES

Français Natif

Anglais C1

Japonais Débutant

À PROPOS

Data Scientist polyvalente, mon champ d'intervention s'étend de l'analyse et de la visualisation des données jusqu'à la mise en production de modèles ML/IA. Je conçois des solutions fiables et orientées métier. Collaborative, autonome et disponible immédiatement.

EXPÉRIENCE PROFESSIONNELLE

DATA SCIENTIST

Juin 2025 - Déc. 2025

Autorité de Régulation des Télécommunications/TIC de CI

CI/FR/ES

Automatisation intelligente & IA | Power Automate, Python, ML supervisé - Déploiement d'une architecture combinant RPA et ML, réduisant de 50% la charge manuelle dans les processus de traitement récurrents.

Architecture MLOps | Python, SQL, Power BI - Mise en place d'un suivi centralisé de +10 modèles en production avec détection automatique de dérives et alertes, améliorant de 40% la réactivité aux anomalies.

Gestion du cycle de vie ML | CI/CD, MLflow - Standardisation du déploiement et des tests automatisés, réduisant de 60% le temps de mise en production.

DATA SCIENTIST

Avril 2024 - Oct. 2024

Institut de la Vision, INSERM

Paris, France

Deep Learning & Fine-Tuning | TensorFlow, CNN, Transformers - Optimisation et ajustement d'architectures CNN et Transformers pour données expérimentales, augmentant de 30% la précision des modèles.

Ingénierie des Données | Python, SQL, Pandas - Création d'une Pipeline reproductible pour +1000 datasets (> 100 Go) avec nettoyage, normalisation, fusion et contrôle qualité complet.

Feature Engineering Avancé | Python, scikit-learn, scipy - Extraction et sélection de 30+ features sur signaux temporels, identifiant 10 biomarqueurs clés pour l'interprétation clinique.

DATA SCIENTIST

Févr. 2021 - Mai 2021

LCQB

Paris, France

Machine Learning Distribué | Python, Dask, Slurm - Optimisation de l'entraînement sur cluster HPC : réduction du surapprentissage et des temps de calcul (-80%).

Ingénierie & Automatisation | Bash, Python, Pandas - Développement de pipelines robustes incluant la gestion d'erreurs pour garantir la reproductibilité des analyses.

DÉVELOPPEUSE FULLSTACK

Mai 2019 - Avril 2020

IKIGAI

Paris, France

Architecture & Ingénierie de Données | Node.js, MongoDB, SQL - Orchestration de pipelines automatisés unifiant 3 millions d'enregistrements hétérogènes pour alimenter un entrepôt de données centralisé.

Développement d'Application Analytique | Angular, TypeScript - Conception d'une interface de visualisation interactive (type Power BI) dédiée à l'exploration de données et au pilotage temps réel (-90% temps d'analyse).

FORMATION

MASTER EN DATA SCIENCE, SPÉCIALISATION EN BIOLOGIE

Sept 2021 - Déc. 2024

Sorbonne Université

LICENCE GÉNÉRALE INFORMATIQUE

Sept 2018 - Sept. 2021

Sorbonne Université

PROJETS PERSONNELS

ANALYSE ENVIRONNEMENTALE SATELLITAIRE

Mai 2025 - Juil. 2025

Détection d'anomalies environnementales par IA sur images satellite.

Précision de segmentation : 91% IoU et 94% Accuracy atteintes après normalisation.

MONITORING GÉOPOLITIQUE MULTILINGUE

Déc 2024 - Févr. 2025

Analyse de tendances et de risques sur données multilingues grâce au Deep Learning.

Fidélité de classification : 95% ROC-AUC et 88% F1 Score obtenus après fine-tuning.

DÉTECTION DE FRAUDE FINANCIÈRE

Mai 2024 - Juil. 2024

Automatisation de l'identification de schémas de fraude.

Performance de détection : 92% AUC sur données déséquilibrées.