

BAHAMOU OMAIMA

Data Scientist · AI Developer · NLP & Machine Learning

+212 654 637 058 · omaimabahamou237@gmail.com · Tanger, Maroc · github.com/omaimabahamou

FORMATION

Master — Bioinformatique & Data Science
ENSA Tanger
2023 – 2025

Licence — Mathématiques & Informatique
Univ. Abdelmalek Saadi, Tétouan
2020 – 2023

CERTIFICATIONS

◆ **Data Analysis with Python**
IBM / Cognitive Class · Avril 2025

◆ **AI Fundamentals**
IBM SkillsBuild · 2024

COMPÉTENCES

▸ Data Science & Analyse

Stats Stat. descriptives/inférentielles, tests d'hypothèse
Feature PCA, sélection, normalisation, encodage, imputation
Validation Cross-validation, GridSearch, AUC, F1, RMSE, MAE
DataViz Matplotlib · Seaborn · Plotly · Power BI · Tableau
Librairies Pandas · NumPy · SciPy · Scikit-learn

▸ Machine Learning / Deep Learning

Algos Régression · Classification · Clustering · Ensembles
DL CNN · RNN · LSTM · Attention · Transformers
Frameworks PyTorch · TensorFlow · Keras

▸ NLP & IA Générative

Modèles BioBERT · FinBERT · HuggingFace · LLMs
Tâches Résumé auto · NER · Classification · Q&A
GenAI Prompt engineering · RAG · Agents IA

▸ Ingénierie & MLOps

Langages Python · SQL · JavaScript · R · PHP
Backend Flask · Django · FastAPI · Node.js
MLOps Docker · Kubernetes · Git · CI/CD
Cloud AWS · Azure · GCP · GPU Infra
Big Data Apache Spark · Snowflake · BigQuery
BDD PostgreSQL · MongoDB · Neo4j · Redis

LANGUES

Arabe Langue maternelle
Français Courant (C1)
Anglais Courant (B2+)

PROFIL

Data Scientist et développeuse IA avec un Master, spécialisée en machine learning, NLP et développement full-stack. Expérience en pipelines de données, modèles prédictifs et IA appliquée (finance, médecine, environnement). Passionnée par l'IA générative, les systèmes agents et la valorisation des données pour des décisions à fort impact.

EXPÉRIENCE PROFESSIONNELLE

Data Scientist — Stage · *Lotus Capital Gestion · Casablanca* 2024 · 2 mois

- Conception d'un scoring automatisé multi-couches pour l'évaluation quantitative de modèles financiers
- Normalisation Z-score avec gestion des outliers, multiplicateurs dynamiques FinBERT (sentiment–position)
- Batch vectorisé pour 100 à 15 000 soumissions — réduction de 60 % du temps d'évaluation manuelle

Digital Marketing — Stage · *Agence ZYRAA* 2023 · 3 mois

- Extraction et nettoyage de données comportementales consommateurs, dashboards interactifs clients
- Modélisation prédictive pour l'optimisation du ciblage et du ROI des campagnes

PROJETS ACADÉMIQUES

Système d'Aide à la Décision Clinique IA 2024

Python · BioBERT · FastAPI · React · ML · ETL

- Modèles ML supervisés pour prédiction de trajectoires patients et outcomes thérapeutiques (AUC > 0.88)
- Modules BioBERT : résumé automatique de notes cliniques et analyse bibliographique en temps réel
- Interface full-stack React + FastAPI, pipeline ETL complet sur données EHR

Drug Discovery Application 2023

Python · BioBERT · Neo4j · ML · PubMed API · RDKit

- Intégration BioBERT + graphes de connaissances Neo4j pour la découverte de candidats médicamenteux
- Prédiction de toxicité hépatique/cardiaque ; recherche PubMed automatisée avec filtrage sémantique
- Visualisation 2D/3D de molécules et analyse de similarité chimique (Tanimoto coefficient)

Big Data Analysis — Spark MLlib 2023

PySpark · K-Means · Régression · Matplotlib

- Segmentation K-Means de personnages de jeux sur features multidimensionnelles (attaque, défense, esport)
- Analyse sismique 1965–présent : régression fréquence-magnitude sur millions d'entrées Spark

Prédiction de Churn Client (Telecom) 2024

Python · Scikit-learn · XGBoost · SHAP · Streamlit

- Modèle de classification binaire (XGBoost, Random Forest) pour prédire l'attrition client — F1 : 0.91
- Feature engineering avancé : variables comportementales, RFM, interactions temporelles
- Interprétabilité avec SHAP values pour identifier les leviers de rétention
- Dashboard Streamlit interactif pour simulation de score de risque par segment client

Analyse de Sentiments & Opinion Mining 2024

Python · Transformers · BERT · NLP · Flask

- Pipeline NLP end-to-end : scraping, nettoyage, tokenisation, analyse de sentiment multi-classes
- Fine-tuning de BERT sur données en arabe et français pour classification d'opinions produits
- API Flask pour scoring en temps réel et visualisation des tendances par thème

Prévision de Séries Temporelles (Énergie) 2024

Python · LSTM · Prophet · Pandas · Plotly

- Prédiction de la consommation électrique nationale avec LSTM et Facebook Prophet
- Détection d'anomalies et décomposition saisonnière pour identification des pics de demande
- Comparaison ARIMA vs LSTM vs Prophet — RMSE réduit de 23 % par rapport au modèle baseline

Système de Recommandation 2023

Python · Collaborative Filtering · SVD · FastAPI

- Moteur de recommandation hybride (filtrage collaboratif + contenu) sur dataset MovieLens 1M
- Décomposition matricielle SVD avec évaluation NDCG@10, Precision@K et recall@K
- API FastAPI pour recommandations en temps réel, déployée sur Docker