

MOHAMED NASSIH

INGÉNIEUR EN DATA SCIENCE



<https://github.com/MohamedNassih>



www.linkedin.com/in/mohamed-nassih



Rabat



medisnas002@gmail.com



+212 6 63 57 24 39

Ingénieur en Data Science récemment diplômé (INSEA), passionné par la résolution de problèmes complexes grâce à l'**analyse de données** et l'**intelligence artificielle**. Rigoureux et autonome, je maîtrise la chaîne de valeur de la donnée, de la **collecte** de données brutes jusqu'au développement de **modèles** prédictifs et à la création de **solutions** interactives. Je cherche maintenant à appliquer ces compétences pour créer des **produits data** à fort impact et contribuer à des projets innovants.



EXPÉRIENCES PROFESSIONNELLES

Data Scientist | Stage PFE | IWRI - UM6P

Février 2025 - Juillet 2025

Développement d'une plateforme Streamlit intégrant nettoyage, agrégation et analyse fréquentielle de séries hydrologiques marocaines. Implémenté des modèles Extreme Value Theory pour estimer les précipitations rares et validé leur robustesse sur trois bassins (Tensift, Draa, Oum Er-Rbia)

Data Scientist | Stage PFA | 3D Smart Factory

[LINK](#)

Aout 2024 - Octobre 2024

Assistant de recherche scientifique avec RAG et embeddings pour répondre aux requêtes des utilisateurs, utilisant des modèles de langage avancés et des représentations sémantiques. Travail sur la segmentation automatique des textes et la génération de réponses personnalisées basées sur le profil utilisateur.

Data Scientist | Stage | YaneCode Digital

[LINK](#)

Aout 2024 - Septembre 2024

Création et évaluation de modèles NLP pour la traduction du darija marocain vers l'anglais. Développement d'algorithmes de collecte de données pour former des modèles. Construction d'un dataset pour des tâches de traduction et d'analyse textuelle en darija.

Stage de découverte | Ministère d'Économie et des Finances - AJR

Juin 2023 - Juillet 2023

PROJETS

Détection du Stress Hydrique Agricole - Bassin du Tensift

[LINK](#)

Septembre 2025

Pipeline complet de Data Science géospatiale pour la détection précoce du stress hydrique agricole par télédétection satellitaire. Combine données Sentinel-2, MODIS et mesures pluviométriques in-situ sur 12 stations du bassin du Tensift

Apprentissage Semi-Supervisé pour les Images Hyperspectrales

[LINK](#)

Décembre 2024 - Février 2025

Développement d'un modèle semi-supervisé combinant données étiquetées et pseudo-étiquetées pour la classification et le mélange spectral d'images hyperspectrales

VPC25 - Voice Privacy Challenge

[LINK](#)

Février 2025

Pipeline d'anonymisation vocale : x-vector pseudo-speaker (ECAPA), DSP (formant warping McAdams, normalisation, pitch-jitter), TTS/Vocoder (optionnel), mapping stable par locuteur, et évaluation (EER/WER, similarité).

Decodage de texte partir de EEG non invasif

[LINK](#)

Novembre 2024 - Janvier 2025

Analyse et modélisation de données EEG pour traduire des états cérébraux simulés en séquences textuelles.

Délimitation multispectrale des Champs Agricoles

[LINK](#)

Décembre 2024 - Janvier 2025

Entraînement d'un modèle UNet pour la délimitation des champs agricoles à partir d'images multispectrales

Évaluation de la pertinence juridique par ML

[LINK](#)

Décembre 2024

Système français Q-A ↔ articles de loi : bi-encodeur CamemBERT pour la similarité, variantes cross-encoder, pipeline d'entraînement/évaluation (MSE, Spearman, F1), jeux de données structurés et notebooks.

COMPÉTENCES

Langages de Programmation

- Python, SQL, R

Analyse de données

- Exploration des données (**Pandas, Numpy**), Réduction de dimension (**ACP**), time series, Data Visualisation (**Matplotlib, Seaborn**), Analyse et traitement géospatial (**GDAL, rasterio, Google Earth Engine**), Extreme Value Theory (**POT, GEV**)

Machine Learning

- (SciKitLearn), Les algorithmes de régression, de classification (SVM, KNN, arbres de décision), d'ensembles (forêts aléatoires, XGBoost), et de clustering (K-means)

Deep Learning & NLP

- (**Tensorflow, Pytorch**), **CNN, RNN** et **LSTM**, (**NLTK, Spacy**), Embedding, Vector database (**Chroma db**), RAG, Transformers (**BERT, GPT, ViT**), Fine-tuning

MLOps & gestion des projets

- Docker, Git, MLflow, Trello, Slack

Soft skills

- Résolution de problèmes, pensée critique, Gestion d'équipe dans des projets data, apprentissage continu et passion pour l'exploration de nouvelles techniques IA

Langues

- Arabe (Maternelle), Français (Courant), Anglais (Intermédiaire)

FORMATIONS

Cycle d'ingénieur, Data Science, INSEA, Rabat, Maroc

2022 - 2025

Institut National de Statistique et d'Économie Appliquée - INSEA

Les Classes Préparatoires aux Grandes Écoles, MPSI - MP, Khouribga

2020 - 2022

Big Bang Data Science Solutions, InnovatiCS, Région métropolitaine d'Atlanta

[LINK](#)

2022 - 2023

CENTRE D'INTÉRÊT



Bénévolat (INSEA CHARITY, GOLDEN MEAN)



Parascolaire (INNOVATION EDGE, INSEA IT)



Technologie



Randonnée