

ABDOULLATUF MAOULIDA

a.ellatuf@gmail.com | 07 82 86 97 64 | Ile-de-France | linkedin.com/in/a-maoulida | abdoullatuf.github.io

FORMATION

Ingenieur en Intelligence Artificielle (RNCP 7) — DataScientest / CentraleSupelec	2024 – 2026
Data Scientist (RNCP 7) — OpenClassrooms	2022 – 2024
Master 2 Probabilites et Statistiques des Nouvelles Donnees — Universite Paris-Est Marne-la-Vallée	2015 – 2016
Master 1 Mathematiques — Universite Paris-Est Creteil	2014 – 2015

COMPETENCES TECHNIQUES

Langages	Python, R, SQL, JavaScript
ML / Deep Learning	scikit-learn, LightGBM, XGBoost, TensorFlow, Keras, HuggingFace Transformers, Optuna
NLP	TF-IDF, Word2Vec, BERT, Universal Sentence Encoder, RoBERTa
Computer Vision	SIFT, VGG-16, ResNet50, InceptionV3 (Transfer Learning)
Data & Viz	pandas, NumPy, Plotly, Matplotlib, Seaborn, SHAP, Power BI
Outils & Cloud	Git, Docker, AWS, Streamlit, PostgreSQL, MongoDB

EXPERIENCE PROFESSIONNELLE

Data Scientist

OpenClassrooms — Formation diplomante certifiee RNCP 7	Sept. 2022 – Fev. 2024
<ul style="list-style-type: none"> Developpe un modele de scoring credit (LightGBM) atteignant un AUC de 0.79 sur 300K+ clients, avec optimisation du seuil de decision via fonction de cout metier personnalisee (cout FN = 10x FP) Concu un systeme de segmentation client non supervise (K-Means + RFM) sur 100K+ commandes Olist, validant la stabilité des segments avec un ARI > 0.72 sur 3 periodes temporelles Implemente un classifieur multimodal (NLP + Vision) benchmarkant 6 approches NLP et 3 architectures CNN, identifiant USE comme meilleur modele texte (ARI 0.49) Deploye 3 dashboards interactifs (Streamlit) integrant des visualisations SHAP pour l'interpretabilite des predictions 	
Statisticien — Stage	
Safran Aircraft Engines — Corbeil-Essonnes	
<ul style="list-style-type: none"> Developpe des methodes d'analyses statistiques multivariees pour l'optimisation du processus d'impression 3D MIM (Metal Injection Molding) en contexte industriel aeronautique 	

Professeur de Mathematiques

Education Nationale (Ac. Versailles & Creteil)	Sept. 2016 – Present
<ul style="list-style-type: none"> Enseigne les mathematiques a 150+ eleves/an au college et lycee, developpant des competences en vulgarisation de concepts complexes, gestion de projets pedagogiques et communication 	

PROJETS

Scoring Credit — « Pret a Depenser »

Python, LightGBM, SHAP, Optuna, Streamlit

- Classification binaire predisant la probabilité de remboursement d'un pret bancaire sur le dataset Home Credit (300K+ clients, 245 features selectionnees)
- Dashboard interactif avec jauge de decision, interpretabilite locale/globale (SHAP waterfall, decision plots), et comparaison client vs. profils moyens

Segmentation Client E-commerce — Olist

Python, K-Means, PCA, Plotly, Streamlit

- Segmentation non supervisee de 100K+ clients via analyse RFM enrichie (Recency, Frequency, Monetary, Satisfaction) avec validation temporelle de la stabilité
- Dashboard avec projection PCA, radar charts, boxplots RFM et comparaison inter-segments pour aide a la decision marketing

Classification Automatique de Produits — Flipkart

Python, BERT, USE, TF-IDF, CNN, Streamlit

- Benchmark de 6 approches NLP (TF-IDF, Word2Vec, BERT, USE, RoBERTa) et 3 CNN (VGG-16, ResNet50, InceptionV3) pour la categorisation automatique de produits e-commerce
- Demo interactive avec prediction top-3 categories et scores de confiance en temps reel

Analyse Statistique Multivariee — Projet STA211 (CNAM)

R, Statistiques multivariees

- Application de methodes d'analyse multivariee avancees (ACP, AFC, classification) sur des jeux de donnees reels dans le cadre du Master Science des Donnees

LANGUES

Francais (courant) • Anglais (intermediaire B1) • Comorien (langue maternelle)