

TABAKH Oussama

Tanger, Maroc | +212 6 41 16 21 00 | eng.oussama.tabakh@gmail.com | LinkedIn: Oussama Tabakh |
GitHub: flash-hero

Profil

Étudiant ingénieur en dernière année, Intelligence Artificielle et Sciences des Données (ENSI). Formation généraliste en data science et machine learning, avec une expérience concrète en IA générative (RAG, LLM/VLM) et extraction de données appliquées à des cas industriels. Recherche un stage de fin d'études (PFE, 6 mois) pour concevoir et déployer des solutions IA à fort impact.

Expérience Professionnelle

Stagiaire Intelligence Artificielle & Développement Backend 07/2026 – 09/2026

Bouygues Construction IT, Casablanca (CasaNearshore) – PFA : moteur d'extraction de données pour l'application eDoc

- Conception d'un moteur d'extraction OCR (Java 17, Spring Boot) structuré en trois traitements indépendants : extraction, classification, validation.
- Développement d'un pipeline hybride associant extraction de texte natif (Apache PDFBox) et lecture par modèle de langage et de vision (LLM/VLM, Azure AI Foundry) pour documents natifs, scannés et grands formats.
- Classification par correspondance floue (distance de Levenshtein) avec garde-fous anti-invention, réduisant nettement la latence et les erreurs de rattachement face à l'alternative testée.
- Intégration REST dans le formulaire de dépôt eDoc, avec préremplissage automatique et suite de tests automatisés (JUnit 5, Mockito, AssertJ).

Data Science & IA Intern

04/2025 – 06/2025

Smart Automation Technologies, Tanger

- Plateforme de classification (Random Forest) avec pipeline complet : prétraitement, sélection de variables, validation croisée.
- Assistant conversationnel basé sur une architecture RAG (LangChain, ChromaDB, LLM), déployé via une API Flask et une interface Streamlit pour utilisateurs non techniques.

Compétences

Machine Learning & Deep Learning : Python, Scikit-Learn, TensorFlow, Keras, CNN, LSTM, RNN

IA Générative & NLP : LLM, VLM, RAG (Retrieval-Augmented Generation), LangChain, LlamaIndex, Prompt Engineering, OCR

Data : Pandas, NumPy, Matplotlib, Plotly, Seaborn, Folium

Bases de Données : MySQL, PostgreSQL, SQL Server, Oracle, PL/SQL, MongoDB, NoSQL, ChromaDB, Talend

Déploiement & Cloud : FastAPI, Flask, Docker, AWS, Azure, Hadoop, Linux, Git/GitHub

Automatisation : n8n, Zapier

Projets Académiques

AI/ML Toolbox – Plateforme de Machine Learning de bout en bout

FastAPI, Pydantic, Scikit-learn, TensorFlow/Keras, Keras Tuner, LightGBM, SciPy

- Développement d'une bibliothèque d'algorithmes de classification, régression et clustering, avec module de deep learning (LSTM/GRU).
- Mise en place d'un pipeline d'EDA et de feature engineering, avec sélection statistique des variables et recherche d'hyperparamètres automatisée.

CV Tailor AI – Système automatisé d'adaptation de CV

Python, NLP, LLM, JSON

- Conception d'un système NLP/LLM adaptant un CV source à une description de poste, avec extraction de mots-clés et scoring de pertinence ATS.
- Génération automatisée de CV personnalisés (DOCX/PDF) en une page, avec sortie structurée en JSON pour limiter les hallucinations du modèle.

Formation Académique

Cycle d'ingénieur – Intelligence Artificielle & Sciences des Données, ENSI, Tanger 2025 – aujourd'hui

Licence Sciences et Techniques – Analyse des Données, FST, Tanger (Mention : Assez Bien) 2021 – 2025

Baccalauréat – Science Physique, option française, Lycée Al Wahda, Oued Amlil (Mention : Bien) 2021

Langues & Certifications

Anglais : EF SET C2 | **Français** : TCF B2 | **Certifications** : Data Scientist Associate – DataCamp (2025), Data Analyst Associate – DataCamp (2025)