

TABAKH Oussama

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

Profile

Final-year engineering student in Artificial Intelligence & Data Science (ENSI). Broad foundation in data science and machine learning, with hands-on experience in generative AI (RAG, LLM/VLM) and data-extraction pipelines applied to industrial use cases. Seeking a 6-month final-year internship (PFE) to design and deploy high-impact AI solutions.

Professional Experience

Artificial Intelligence & Backend Development Intern 07/2026 – 09/2026
Bouygues Construction IT, Casablanca (CasaNearshore) – Capstone project: automated data-extraction engine for the eDoc document management app

- Designed an OCR extraction engine (Java 17, Spring Boot) structured into three independent stages: extraction, classification, validation.
- Built a hybrid pipeline combining native text extraction (Apache PDFBox) with language/vision model reading (LLM/VLM, Azure AI Foundry) for native, scanned, and large-format documents.
- Implemented fuzzy-matching classification (Levenshtein distance) with anti-hallucination safeguards, significantly cutting latency and mismatch errors versus the tested alternative.
- Integrated the engine via REST into the eDoc submission form, with automatic field pre-filling and an automated test suite (JUnit 5, Mockito, AssertJ).

Data Science & AI Intern 04/2025 – 06/2025
Smart Automation Technologies, Tangier

- Built a classification platform (Random Forest) with a full pipeline: preprocessing, feature selection, cross-validation.
- Developed a RAG-based conversational assistant (LangChain, ChromaDB, LLM), deployed via a Flask API and a Streamlit interface for non-technical users.

Skills

Machine Learning & Deep Learning: Python, Scikit-Learn, TensorFlow, Keras, CNN, LSTM, RNN
Generative AI & NLP: LLM, VLM, RAG (Retrieval-Augmented Generation), LangChain, LlamaIndex, Prompt Engineering, OCR
Data: Pandas, NumPy, Matplotlib, Plotly, Seaborn, Folium
Databases: MySQL, PostgreSQL, SQL Server, Oracle, PL/SQL, MongoDB, NoSQL, ChromaDB, Talend
Deployment & Cloud: FastAPI, Flask, Docker, AWS, Azure, Hadoop, Linux, Git/GitHub
Automation: n8n, Zapier

Academic Projects

AI/ML Toolbox – End-to-End Machine Learning Platform
FastAPI, Pydantic, Scikit-learn, TensorFlow/Keras, Keras Tuner, LightGBM, SciPy

- Built a library of classification, regression, and clustering algorithms, with a deep learning module (LSTM/GRU).
- Set up an EDA and feature-engineering pipeline, with statistical feature selection and automated hyperparameter search.

CV Tailor AI – Automated Resume-Tailoring System
Python, NLP, LLM, JSON

- Designed an NLP/LLM system that tailors a source resume to a job description, with keyword extraction and ATS relevance scoring.
- Automated generation of one-page tailored resumes (DOCX/PDF), with structured JSON output to limit model hallucinations.

Education

Engineering Degree – Artificial Intelligence & Data Science, ENSI, Tangier 2025 – Present
Bachelor's Degree (Licence) – Data Analysis, FST, Tangier (Honors: Assez Bien) 2021 – 2025
Baccalaureate – Physical Science, French track, Lycée Al Wahda, Oued Amlil (Honors: Bien) 2021

Languages & Certifications

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