

Mouaad El Yalaoui

+212 719-112489 | Rabat | moad2632005@gmail.com | github.com/Moad26 | linkedin.com/in/mouaad-el-yalaoui

EDUCATION

Institut National des Postes et Télécommunications Rabat
Data science Engineering 2024 — 2027

CPGE Moulay Ali Cherif Errachidia Errachidia
National Ranking: 127th 2022 — 2024

WORK EXPERIENCE

Research assistant intern June 2025 — August 2025

Institut National des Postes et Télécommunications

- Engineered a Kubernetes failure prediction system using Bi-LSTM (with attention), Isolation Forest, and SVM.
- Implemented multi-task learning for failure detection and severity classification across 77+ temporal features.
- Optimized models via grid search with custom scoring and generated 27+ Plotly visualizations for monitoring.

PROJECTS

Open-Books: Local RAG System for Technical PDFs | Python, ChromaDB, Ollama, Docling, sentence-transformers, Textual, FastAPI, Redis

- Engineered local RAG system for technical PDFs using **Docling** (extraction), **Sentence-Transformers**, and **ChromaDB**.
- Eliminated redundant file processing by implementing hash-based incremental indexing and caching.
- Built **Textual** TUI for real-time semantic search and local inference via **Ollama**.
- Architected modular pipeline with section-aware text splitting and configurable chunking strategies.

Music Recommender System via Audio Signal Processing | Python, librosa, scikit-learn, XGBoost, FastAPI, Streamlit, Docker

- Built recommender analyzing **1.2M tracks** via dual feature extraction (MFCC/statistical) and cosine similarity.
- Achieved **0.68 F1** (mode) and **0.45 R²** (energy) using a hybrid stacking ensemble (RF, XGBoost, SVM).
- Compressed 470+ features to 155 components via **PCA**, retaining **>90% explained variance**.
- Deployed via **FastAPI**, **Streamlit**, and **Docker** with intelligent spectrogram caching.

Transformer for Empathetic Dialogue (From-Scratch Implementation) | PyTorch, Transformers, HuggingFace, pandas, NumPy, Architecture Design

- Implemented custom **PyTorch** encoder-decoder Transformer from scratch on **EmpatheticDialogues** (25k conversations).
- Engineered multi-head attention with independent modules per head for enhanced representation capture.
- Designed training pipeline with **Adam** optimization, masked cross-entropy, and early stopping.
- Built preprocessing pipeline using **BlenderBot** tokenization with dynamic padding and truncation.

EXTRACURRICULAR ACTIVITIES

MSC Research cell Oct 2025 — Present

- Co-authored a research paper on Music Emotion Recognition.

Won the CIT's Ramadan League Data Science Competition 2024

SKILLS

Programming Languages: Python, JavaScript, C/C++, Java, Bash, R, HTML/CSS, Flutter, Dart, Golang, SQL

ML/AI Frameworks & Libraries: PyTorch, TensorFlow, scikit-learn, XGBoost, Transformers, NumPy, pandas

LLM & RAG Technologies: Ollama, ChromaDB, Redis, RedisVL, Opik (ML Tracing), Logging (Loguru)

Web Development: React, FastAPI, Streamlit, Pydantic, Htmx.

DevOps & Infrastructure: Docker, Docker Compose, Git, GitHub Actions, CI/CD, Kubernetes.

Languages: English (fluent), French (fluent), Arabic (native)