

Hamza Boulahia

hamzaboulahia@gmail.com | [in /hamzaboulahia](https://www.linkedin.com/in/hamzaboulahia) | [/Hmzbo](https://github.com/Hmzbo) | [/Hmzb.github.io](https://hmzb.github.io)



WORK EXPERIENCE

Freelancer

(May 2024 – Present)

Freelance AI Engineer - AI Inheritance Chatbot

- Led dataset creation and preprocessing for domain-specific Islamic inheritance Q&A, preparing data for both fine-tuning and RAG.
- Experimented with quantized LoRA fine-tuning of open-source LLMs using augmented Arabic data; evaluated via ROUGE, BLEU and custom LLM-based metrics, and concluded RAG was preferable.
- Benchmarked Arabic embedding models to select the best for vector indexing.
- Designed, implemented, and evaluated multiple RAG pipelines (simple RAG, reranking, LLM-judged), using classical NLP, custom LLM, and RAG-specific metrics to identify the optimal setup.
- Integrated the chosen RAG retrieval module into the chatbot to deliver accurate, context-aware inheritance answers in Arabic.
 - **Tech stack:** Python, LangChain, FAISS, Hugging Face Transformers, PyTorch, unsloth, ragas, Gradio.

Data Scientist

Aléria Consulting (May 2022 – May 2024)

- Developed a web application for webpage attention heatmap prediction
 - Training and testing a deep learning model to predict webpage saliency.
 - Achieving more than 85% accuracy on the Webpage Saliency public dataset.
 - Implemented a robust Python API for seamless model integration.
 - **Tech stack:** Python, U-Net, OpenCV, Flask
- Designed a Windows desktop application for ID Optical Character Recognition (OCR)
 - Creating and optimizing object detection models for efficient extraction of information from ID cards.
 - Applied advanced image processing techniques to minimize OCR errors.
 - Achieving accurate and reliable extraction of data from ID cards, and delivering product within due time.
 - **Tech stack:** Python, YOLO, OpenCV, Tesseract-OCR, PyQt5

Freelancer

(Oct 2020 – May 2022)

Data Science Consultant

- Creating and improving classification and regression machine learning models using Python as per client's requirements.
- Performing exploratory data analysis to extract insights and delivering detailed reports.

Math Tutor

Sigma Training Center (Oct 2016 – Aug 2018)

- Tutoring students in bachelor level mathematics (Calculus, Linear Algebra, and Convex Optimization).
- Incorporating visuals using Python to help students better understand certain math concepts.

EDUCATION

Research Master in Mathematical Modeling and Scientific Calculations

Relevant subjects: Statistics, Stochastics, Optimization

National Engineers School of Tunisia (2014 – 2016)

Bachelor of Science in Applied Mathematics

Relevant subjects: Linear Algebra, Calculus, Statistics, Algorithmics & Data Structures

Higher School of Science and Technology of Tunis (2011 – 2014)

CERTIFICATIONS

- TensorFlow 2 for Deep Learning ([Coursera](https://www.coursera.org/learn/tensorflow-2-for-deep-learning) - Aug 2021)
- Deep Learning Specialization ([Deeplearning.ai](https://www.deeplearning.ai) - Oct 2020)
- Machine Learning ([Coursera](https://www.coursera.org/learn/machine-learning) – Apr 2020)
- Data Analyst Nanodegree ([Udacity](https://www.udacity.com) | Oct 2019 – Jun 2020)

SKILLS

Languages: Arabic (Native), French (Professional), English (Professional).

Programming & ML Frameworks: Python (Pandas, PyTorch, TensorFlow, FastAPI, Flask, Plotly, Streamlit, Gradio), SQL.

NLP & Computer Vision: LLM Fine-tuning, RAG, AI Agents (LangChain, LangGraph, LlamaIndex), BERT, spaCy, OpenCV, YOLO models, SAM models.

Tools & DevOps: Git, CI/CD (GitHub Actions), AWS, Docker, MySQL, PyQT5, Selenium

PROJECTS & CONTENT CREATION

- **YouTube Channel: [Machine Learning With Hamza](#)**
 - Creator & Host of ~32 educational videos focusing on machine learning and AI tutorials; channel grew to ~5.5K subscribers as of June 2025
 - Produced content on topics such as neural networks, fine-tuning LLMs, chatbots, NLP and CV techniques, and Reinforcement Learning.
 - Managed end-to-end production: topic research, scripting, recording, editing, and thumbnail design; solicited and responded to viewer questions, fostering community engagement and feedback loops.
- **Web application for football video analysis** ([Github](#) | YOLOv8, OpenCV, Streamlit):
Creating a web application for football object detection with team prediction and tactical map representation.
- **Customized chatbot web application** ([Github](#) | GPT-3.5, Gradio, Tacotron 2):
Creating a chatbot web application for lyrics generation and speech synthesis using deep learning model for voice cloning.
- **Web scraper for job offers on LinkedIn** ([Github](#) | Python, Selenium):
Creating a web scraper to collect job offers data based on a job title and location, and save the extracted data in csv files.

INTERESTS

Reading, educational videos creation, and calisthenics workout.