Mohamed Alawakey

mohamedelawakey@gmail.com | +201127247680 | Cairo, Egypt | mohamedelawakey.com | linkedin.com/in/mohamedelawakey

github.com/mohamedelawakey

Education

Bachelor of Computers and Informatics with a minor in Information Systems, *Tanta University* — 10/2022 – 08/2026, Tanta, Egypt
Relevant Coursework: Machine Learning, Deep Learning, NLP

Professional Experience

GenAI Engineer, *Ghaymah Systems* — 09/2025 – Present, Remote
- **Developed MCP servers** from scratch to integrate **LLM-powered** applications with **GitHub** and **GitPasha**
- **Optimized Retrieval-Augmented Generation (RAG)** pipelines by **implementing semantic reranking** and **chunk optimization**, **improving retrieval accuracy** by **20%**

Data Scientist Trainee, *ALX Africa* — 01/2025 – 03/2026, Hybrid, Egypt
- **Developed** and **evaluated 5+** practical **machine learning** projects using **Python** and **SQL** for end-to-end **data processing**, achieving an **85-95% average accuracy** across standard **evaluation metrics**

Artificial Intelligence Intern, *NTI* — 04/2025 – 05/2025, Tanta, Egypt
- **Completed** intensive **AI**, **ML**, and **NLP** training with a **97% graduation score**, **developing** and **fine-tuning** a capstone **machine learning model** that achieved **92% evaluation accuracy**

Projects

QualiVerse (Deployed at Tanta University)
- **Led a cross-functional team of 9 members** from initial **architectural planning** to full-scale **deployment**, delivering the winning system among **5** competing candidates for official implementation at the Faculty of Computers and Information, Tanta University
- **Architected a high-performance microservice using FastAPI and LLMs**, seamlessly integrated with a larger .NET ecosystem, and deployed via **Hugging Face** to automate academic documentation with **85-92% content accuracy**
- **Engineered an AI-driven Course Specification pipeline** using **Prompt Engineering**, slashing manual processing time by **75-90%** (**reducing** a **3-hour** manual task to just **5-10 minutes**)
- **Developed an automated Course Report generator** that optimized reporting workflows by **80-90%**, reducing the turnaround time from **2 hours** of manual effort to under **3 minutes** per report

RAGU (Comprehensive Educational AI Platform)
- **Architected** an **end-to-end** educational **RAG** platform for daily academic study using **Milvus**, **PostgreSQL**, and **Sentence-Transformers**, achieving **1.0 Faithfulness**, **1.0 Context Recall**, and **93.6% Answer Relevancy** via **Ragas evaluation**
- **Applied Prompt Engineering** and **Cost Optimization** to build secure, tier-based prompts with injection defenses, **slashing token consumption by 60-65% (e.g., from ~25k to ~5k)** across document Q&A, mind-maps, and automated quiz workflows
- **Integrated a PaddleOCR pipeline**, **reducing** Docker footprint and memory usage by **20% via OpenCV** dependency **optimization**, and utilized LLM-based image extraction to cut Vision API costs by **70-80%**
- **Developed scalable FastAPI** endpoints for **hybrid retrieval** models and tested a **local MVP** via **Cloudflare Tunnels**, securing **25+ positive functional validations** from university professors and peers

AI Job Hunter System Agent (Open Source)
[Github]
- **Developed** an **Agentic AI** assistant using **LangGraph** and **LangChain**, integrating **web scraping** and **LLM** analysis to automate job filtering and extract key requirements
- **Built** a local PDF/DOCX CV-processing pipeline using **Ollama** and **Pydantic** for structured data extraction and validation

Programming Book Recommendation System (Open Source)
[Github]
- **Built** an AI-powered **semantic search engine** using **PyTorch** and **Sentence-Transformers**, reducing a **+10M** dataset to **+400K** records via **Pandas** and **Scikit-learn feature engineering**
- **Developed** a scalable **Docker**-containerized **FastAPI** backend with **Redis** and **PostgreSQL** (**PGVector**/**HNSW**) for **sub-100ms latency**, fully automated via **GitHub Actions CI pipelines**

Student Performance Predictor (Open Source)
[Github] | [Live Demo]
- **Built** a **machine learning** model using **Scikit-learn** to predict student performance, achieving **91% accuracy**
- **Processed** and **analyzed** data using **Pandas** and **NumPy**, including **feature engineering** and exploratory **data analysis**
- **Developed** and deployed a **FastAPI API** integrated with the **ML** pipeline for real-time predictions
- **Containerized** the application using **Docker** and **deployed** it on **Hugging Face**

Skills

Languages & Frameworks — Python, Scikit-learn, PyTorch, NumPy, Pandas, Matplotlib, Seaborn, FastAPI, LangGraph, LangChain

Machine Learning & AI — Deep Learning, LLMs, Prompt Engineering, Data Cleaning, EDA, Data Preprocessing

Tools — Hugging Face, PostgreSQL, Vector Databases, Git, GitHub, Docker, Docker Compose, GitHub Actions, CI/CD

Languages

Arabic (Native) | English (B2)