

Yousef Elgohary

AI Engineer (NLP & LLMs)

yousefelgohary.ai@gmail.com | +201022682019 | Cairo, Egypt (Willing to Relocate)

linkedin.com/in/yousefelgohary | github.com/yousefelgohary

Summary

AI Engineer with **over 2 years of applied experience** building **end-to-end AI systems** through intensive internships and hands-on projects. Specialized in Machine Learning, NLP, **Large Language Models (LLMs)**, **advanced RAG pipelines**, and **Agentic workflows** using **Python, FastAPI, LangChain, and LangGraph**. Proactive engineer with a demonstrated ability to **lead technical teams** and deliver scalable AI solutions.

Education

Bachelor of Electrical Engineering in Communications and Computer Engineering 09/2021 – 09/2026
Higher Institute of Engineering and Technology in Tanta

Experience

Machine Learning for Data Analysis Intern 07/2026 – Present
National Telecommunication Institute (NTI)

- Engineered **ApexML**, an automated, end-to-end machine learning platform deployed on **Streamlit** for dynamic data profiling and model training. Integrated the **Gemini API** to generate intelligent, AI-driven recommendations for optimal data preprocessing and feature engineering.
- Architected and deployed 15+ interactive machine learning web applications utilizing **Supervised and Unsupervised Learning** algorithms (SVM, Neural Networks, Decision Trees, K-Means, PCA).
- Designed scalable data analysis pipelines utilizing **Pandas and NumPy** for end-to-end Exploratory Data Analysis (EDA), applying advanced data cleaning and feature engineering techniques to extract actionable insights.

Natural Language Processing Intern 08/2025 – 09/2025
National Telecommunication Institute (NTI)

- Engineered **Avokat AI**, a bilingual legal assistant using **Gemini LLMs, RAG, and Neo4j Knowledge Graphs** to analyze contracts, extract legal loopholes, and provide grounded document analysis in Arabic and English.
- Developed **EmpathyBot**, a sentiment-driven chatbot utilizing **DistilBERT** and **FAISS-powered RAG**, achieving over **80% accuracy** in real-time emotion detection and delivering context-aware responses.
- Architected backend logic for AI models by building robust **RESTful APIs using FastAPI**, implementing advanced prompt engineering, and optimizing text extraction pipelines using **PyMuPDF**.

Artificial Intelligence Intern 07/2024 – 08/2024
Telecom Egypt

- Co-developed and optimized** Machine Learning and Deep Learning models using **Python** to address data-driven challenges.
- Automated** data preprocessing and model training pipelines by developing custom Python scripts, significantly reducing manual effort.

Projects

Rashaad 09/2025 – Present
A Multi-tenant SaaS AI Academic & Administrative Advisor for Higher Education Institutions.

- Leading and orchestrating a cross-functional team of 8 members**, managing the project's development lifecycle and ensuring seamless collaboration between AI and Web development tracks.
- Architected and implemented an Advanced RAG Fusion system** using **LangChain** for orchestration and **PostgreSQL (Pgvector)** for high-performance vector storage, ensuring high-precision, **bilingual (Arabic/English)** retrieval.
- Engineered the technical bridge** between the **FastAPI** AI service and the **Laravel-based** business logic to support a scalable, decoupled **microservices architecture**.

A zero-hallucination, source-cited AI co-pilot for Gastric Cancer clinical decision support.

- **Architected** an 8-node state machine using **LangGraph** to build an **Agentic RAG** system, providing strictly grounded medical answers based on WHO and ESMO guidelines.
- **Engineered** a sequential AI pipeline leveraging hybrid retrieval (pgvector + BM25 with RRF) for precise context matching, Gemini for **grounded generation**, and an **NLI faithfulness checker** to strip contradicted claims.
- **Developed** a scalable FastAPI backend with SSE streaming, implementing **LangSmith** for full pipeline tracing and **RAGAS** for rigorous evaluation of faithfulness and context precision.

Competitions & Awards

4th Place Winner - AI Hackathon 2026 (Menoufia)

08/2026

Orange Digital Center, ITIDA, TIEC & Instant Software Solutions

- **Led the 5-member** "AI Innovators" team, securing **4th place** in Menoufia Governorate at a nationwide hackathon featuring **2,000+ participants** from 17 governorates in Egypt.
- Successfully built and pitched "**Sadeed**," a **zero-hallucination AI co-pilot** for **Clinical Decision Support (CDS)**.

Courses

[AWS: Becoming a Machine Learning Engineer](#) — Manara (2026)

Getting Started with Deep Learning — NVIDIA (2025)

Python: Programming Fundamentals — Almdrasa (2024)

Skills

Technical Skills:

- **AI & Machine Learning:** Large Language Models (LLMs), RAG & Agentic Workflows, LangChain, LangGraph, Hugging Face, NLP (SpaCy, NLTK), Scikit-learn, Pandas, NumPy.
- **Backend & Databases:** Python, FastAPI, PostgreSQL (pgvector), Neo4j (Knowledge Graphs), FAISS, ChromaDB, RESTful APIs.
- **Development Tools:** Git, GitHub, Linux, Docker, n8n, Streamlit.

Project Management Tools:

- Jira, Notion.

Languages:

- **Arabic:** Native
- **English:** Intermediate (B1)