

Omar Hafez Khalil

Artificial Intelligence Engineer

✉ mo5rb600@gmail.com | 📞 +974 5007 2807 | 📍 Doha, Qatar

🌐 linkedin.com/in/omar-khalil-55a674281

🌐 portfolio-omar-nu.vercel.app | 🐙 github.com/OmarHKhalil

Professional Summary

AI Engineer specializing in Machine Learning, Deep Learning, Computer Vision, and Generative AI, with hands-on experience designing and deploying end-to-end AI systems. Experienced in developing production-oriented Computer Vision pipelines, Retrieval-Augmented Generation (RAG) applications, and multi-agent AI workflows. Skilled in PyTorch, FastAPI, LangChain, LangGraph, LLMs, vector databases, hybrid retrieval, and local LLM deployment, with a strong focus on building practical, scalable, and deployable AI solutions.

Experience

Freelance AI Engineer May 2025 – Present Doha, Qatar

- Engineered end-to-end AI systems spanning Computer Vision, RAG applications, and multi-agent workflows using PyTorch, FastAPI, LangChain, and LangGraph.
- Developed and deployed production-ready microservices for banking systems, document intelligence, and multilingual receipt recognition with high-throughput processing.
- Integrated object detection, OCR engines, vector databases (ChromaDB), and local/cloud LLM inference (Ollama, Gemini API) into scalable pipelines.

Technical Skills

Programming Languages: Python | JavaScript | SQL | C++ | Java | HTML& CSS

Machine Learning & Deep Learning: PyTorch | Scikit-learn | Neural Networks | Transfer Learning | Fine-Tuning | Hyperparameter Tuning | Model Evaluation | Feature Engineering

Computer Vision: OpenCV | Object Detection | Image Classification | OCR | Ultralytics (YOLOv8) | ResNet-50

Generative AI & LLM Engineering: LLMs | RAG | AI Agents | Model Context Protocol (MCP) | LangChain | LangGraph | Transformers | Embeddings | Prompt Engineering | Vector Databases (ChromaDB) | LLM APIs & Local Inference (Gemini, ChatGPT, Grok, Ollama) | AI Security | Prompt Guardrails | Access Control

Model Serving & APIs: FastAPI | REST APIs | AI inference endpoints | Asynchronous APIs

Backend & Databases: Node.js | Express.js | MongoDB | Mongoose | Database Design

Developer Tools: Git | GitHub | Docker | Postman | Jupyter Notebook | Google Colab

Projects

Venezia Bank – AI-Powered Digital Banking System | AI, Flutter, Laravel, MySQL, FastAPI

- Developed a comprehensive AI-powered digital banking ecosystem featuring a cross-platform mobile application for customers and a web-based platform for bank employees, supported by a Laravel backend and integrated AI services. The platform provides core banking services including payments, transfers, account management, loans, investments, financial reporting, and fraud monitoring.
- Served as an AI Engineer designing and implementing 4 core AI modules, deployed as FastAPI microservices to power a comprehensive ecosystem integrating 8+ AI capabilities (including financial fraud detection, receipt recognition, customer churn prediction, product recommendation, stock prediction, loan and employee decision support, and an LLM/RAG chatbot).
- 🐙 **GitHub Repository**

Receipt Recognition System | FastAPI, Ultralytics(YOLOv8), PyTorch, ResNet-50, Gemini API, OpenCV

- Developed an end-to-end AI pipeline converting multilingual receipt images into structured financial data through 4 automated stages: validation, region detection, OCR, and product classification.
- Achieved 98.83% validation accuracy, 0.7693 mAP@50, and 0.93 F1-score across 15 product categories.
- Deployed the pipeline as a FastAPI REST service with 12 asynchronous workers, supporting Arabic and English receipts.

-  **GitHub Repository**

DocMind AI (Multi-Agent) | *Python, LangChain, LangGraph, ChromaDB, Hybrid Retrieval, Ollama, Streamlit*

- Built a custom multi-agent document intelligence pipeline orchestrating 4 specialized agents for relevance checking, extraction, reasoning, and response generation.
- Implemented hybrid retrieval, persistent ChromaDB storage, and document caching for reusable local knowledge.
- Developed an offline-first document analysis application supporting PDF, DOCX, TXT, and Markdown using local Ollama LLMs.
-  **GitHub Repository**

Banking Finance Chatbot (RAG) | *FastAPI, LangChain, ChromaDB, Ollama, Gemini API, Embeddings*

- Developed a domain-specific RAG chatbot for banking and financial assistance using vector retrieval and LLM-based response generation.
- Integrated ChromaDB, embeddings, and local/cloud LLM inference through Ollama and Gemini API.
- Achieved 88.29% BERTScore in response evaluation and deployed the application as a FastAPI service.
-  **GitHub Repository**

AI-powered Banking Retention & Product Recommendation System | *FastAPI, XGBoost, LightGBM*

- Developed an intelligent banking retention system that predicts customer churn and automatically recommends personalized financial products through a real-time REST API. Implemented a dual-machine-learning pipeline for prediction and recommendation.
- Achieved 0.88 F1-score for both churn prediction and product recommendation.
-  **GitHub Repository** |  **Live Demo**

NutriVision: Food Classification | *PyTorch, ResNet50, EfficientNet-B0, TensorBoard, TorchMetrics*

- Developed a deep learning food classification system covering 30 categories, comparing custom CNN and transfer learning architectures.
- Evaluated Custom CNN, ResNet-50, and EfficientNet-B0 models and deployed an interactive inference application.
-  **GitHub Repository** |  **Live Demo**

Publications

Co-Author, “Dual Model Framework for Stock Price Movement Prediction Using Technical Indicators and Financial News Sentiment with Decision Level Fusion”

8th IEEE Jordan Conference on Applied Electrical Engineering and Computing Technologies (AEECT 2026) July 2026

- Research conducted as part of the Venezia Bank AI-powered digital banking system, focusing on a decision-level fusion framework combining technical indicators and financial news sentiment for stock price movement prediction.

Education

Arab International University (AIU)

2021 – 2026

- Bachelor’s Degree in Informatics & Communications Engineering (Specialization: Artificial Intelligence)
- GPA: 3.36/4.0

Courses

- **Agentic AI with LangChain and LangGraph** (Coursera, Completed Jun. 2026)
- **Machine Learning** (Coursera, Completed Mar. 2026)
- **Generative AI with Large Language Models** (Coursera, Completed Sep. 2025)
- **Deep Learning & Computer Vision**
 - **PyTorch for Deep Learning & Machine Learning** (freeCodeCamp, Jan. 2025 – Aug. 2025)
- **Machine Learning with Python and Scikit-Learn** (freeCodeCamp, Jul. 2024 – Dec. 2024)

Languages

Arabic: Native · **English:** Professional Working Proficiency