

HAMMAD FAROOQ

Data Scientist | AI & ML Engineer

Lahore, Pakistan | +92 3450710970 | hammadfarooq470@gmail.com | [linkedin.com/in/hammadfarooq-ai](https://www.linkedin.com/in/hammadfarooq-ai) | github.com/hammadfarooq-ai | kaggle.com/hammadfarooq470

PROFESSIONAL SUMMARY

Results-driven Data Scientist and AI Engineer with hands-on experience building production-grade AI applications, LLM integrations, and end-to-end ML pipelines. Proficient in Python, LangChain, PyTorch, and Azure AI services. Proven track record developing RAG-based chatbots, AI voice assistants, and forecasting systems. Adept at architecting scalable, cloud-ready AI solutions and integrating large language models into real-world workflows.

CORE TECHNICAL SKILLS

Languages & Frameworks: Python, FastAPI, Flask, REST APIs, Object-Oriented Design
LLMs & AI: LangChain, OpenAI GPT-4o, Ollama, RAG Pipelines, Agents, Memory, LangServe
Machine Learning: Scikit-learn, PyTorch, ARIMA, Random Forest, Feature Engineering
Cloud & DevOps: Azure AI Services, Azure ML, Azure OpenAI, Docker, Git, GitHub, CI/CD
Data & NLP: Pandas, NumPy, FAISS, Chroma, Deepgram, ElevenLabs, Speech Recognition
Databases: PostgreSQL, MySQL, FAISS, Chroma (Vector Databases)

PROFESSIONAL EXPERIENCE

Freelance Data Scientist & AI Developer | Self-Employed | Remote

Feb 2025 – Present

- Architect and deliver end-to-end AI solutions using Python, LangChain, LLMs (GPT-4o, Ollama), and FastAPI for international clients.
- Build production RAG pipelines integrating vector databases (FAISS, Chroma) with multi-source document ingestion and retrieval.
- Develop AI-powered agents with memory, tool use, and multi-step reasoning using LangChain Agents and Chains.
- Deploy scalable AI applications on cloud platforms with FastAPI and LangServe; manage full CI/CD workflows via GitHub.

Kaggle Grandmaster | Data Scientist

Apr 2023 – Present

- Achieved **2x Kaggle Grandmaster** ranking in Datasets (#14 globally) and Notebooks (#18 globally).
- Published 620+ notebooks and 154+ datasets on machine learning, NLP, and AI systems.
- Ranked **32/2673 teams** in the Akkadian Translation competition using fine-tuned ByT5 models.
- Built ML pipelines, RAG systems, predictive models, and hybrid retrieval solutions using Python and Scikit-learn.
- Active participant in global AI competitions including ARC-AGI, Legal Information Retrieval, and NVIDIA reasoning challenges.

Machine Learning Engineer Intern | Arch Technologies | Islamabad (Remote)

Apr 2025 – May 2025

- Built ARIMA time-series forecasting models with temporal feature engineering for business analytics.
- Developed regression pipelines for house price prediction using Scikit-learn with end-to-end preprocessing.
- Implemented CNN-based image classification on MNIST using PyTorch for handwritten digit recognition.

KEY AI & PYTHON PROJECTS

LangChain RAG + Agents Chatbot — Python, LangChain, OpenAI, FastAPI, FAISS, Chroma

- Built production-grade LLM workflows using LangChain Chains, Agents, Memory, and RAG with OpenAI and Ollama; deployed with FastAPI and LangServe as a scalable API-first AI service.
- Implemented embedding pipelines, vector databases, and multi-source data integration for intelligent document Q&A.

AI Voice Legal Intake Assistant — Python, LiveKit, OpenAI GPT-4o, Deepgram, ElevenLabs, PostgreSQL, Twilio

- Developed a multi-agent AI voice assistant for automating legal client intake through real-time phone calls using LLMs and speech AI.
- Implemented speech recognition, structured case intake workflows, and intelligent data extraction with Flask-based monitoring APIs.
- Integrated Twilio SIP, SendGrid, PostgreSQL, and Zapier for automated call routing, intake reporting, and CRM workflow automation.

Dogs vs Cats Image Classification — Deep Learning Project

- Built a CNN-based image classification model using Transfer Learning (MobileNetV2) to classify dogs and cats.
- Applied data preprocessing and augmentation techniques to improve model generalization and performance.
- Achieved ~96% validation accuracy and evaluated model using confusion matrix, ROC-AUC, and Grad-CAM visualization.

EDUCATION & CERTIFICATIONS

Bachelor of Science in Data Science — Virtual University of Pakistan

Intermediate in Computer Science (ICS) — Punjab Group of Colleges

Certifications: IBM AI Engineering Professional (Aug 2025) | Microsoft AI & ML Engineering (Jul 2025) | Google Advanced Data Analytics (Nov 2025) | Statistics with Python – Univ. of Michigan (Jul 2025)

ACHIEVEMENTS

Kaggle Top 10% & 2(x) Grandmaster | IBM & Microsoft Certified AI/ML Engineer | LabLab.ai Hackathon Competitor