

MUHAMMAD USMAN KHALIL

AI ENGINEER / WEB DEVELOPER

Lahore, Pakistan | usman.data002@gmail.com | +923098906870

Experienced in artificial intelligence, machine learning, deep learning, neural networks, agentic AI, and LLM fine-tuning using Python, TensorFlow, and Keras. Proficient in designing, deploying, and maintaining production-ready RESTful and scalable API solutions using React, Django, and FastAPI. Skilled in developing intelligent systems including multi-agent AI architectures, RAG pipelines, fine-tuned LLMs, and voice-based AI agents using LangGraph for real-time, context-aware interactions. Strong background in full-stack software engineering with hands-on experience in PHP, SQL, Django, and React. Experienced in Power BI, data analytics, and advanced data visualization to translate complex data into actionable insights and support data-driven decision-making.

AREA OF EXPERTISE

- Machine Learning (tensorflow, pytorch)
- Neural Network(keras)
- LLM, RAG, VectorDB (pinecone, chroma)
- Multi agentic AI, n8n
- LangChain, crewai
- MCP (Model Context Protocol)
- Cloud Deployment (AWS, Azure)
- Web development (django, PHP, React, FastAPI)
- Github
- Data Analyst (Excel, python, PowerBI)

PROFESSIONAL EXPERIENCE

- One Machine Software** July 2025 - Present

Python/ML developer

Developed a Retrieval-Augmented Generation (RAG) pipeline using Pinecone for efficient medical data retrieval and analysis. Implemented and fine-tuned a Large Language Model (LLM) using LSTM architecture for domain-specific text understanding and prediction.

- University Of Lahore (SGD campus)**

Database Intern

Developed and implemented an organizational database to oversee student assessments, including quizzes and assignments, which resulted in improved visibility of key performance metrics and enabled easier decision-making in grading at the organizational level.

PROJECTS

Plant Disease Detection and Treatment

- The project uses a Convolutional Neural Network (CNN) to analyze plant images and determine if they are healthy or diseased and a RAG technique to get the treatment.
- A FastAPI-based backend handles image upload, prediction, and model inference efficiently.

Liver Cancer Detection and Diagnosis using Deep Learning (Research Project)

- Developed a binary classification system using EfficientNet-B0, TinyViT, and MobileViTv2 models to detect Hepatocellular Carcinoma (HCC) from DICOM CT scan images converted to JPG format, achieving up to 98%.
- Deployed the trained models using FastAPI, enabling image upload, inference, and HCC prediction through a user-accessible web API.

RAG Chatbot for Health Insurance Policies

- Built AI automation workflows using CrewAI, FastAPI, and RAG pipelines to enhance operational efficiency and contextual intelligence.
- Optimized and deployed scalable multi-agent applications using FastAPI, Azure, Docker, asynchronous processing, and multithreading.

EDUCATION

Bachelor of Science in Computer Science

University of Lahore

- FYP on "Technological Advancements within the current Agriculture Industry" using ML techniques and RAG.

ADDITIONAL INFORMATION

- Languages:** English, Urdu.
- Certifications:** Probability & Stat for ML (coursera), Ethical Hacker Course (Cisco), Data Analytics & AI workshop (atomcamp).
- Awards/Activities:** First Position in speed programming competition(2024), Onboarding Project Lead in internship.

LINKS



www.linkedin.com/in/khalilusman70



<https://github.com/khalilusman>