Muhammad Ahsan

AI/ML Engineer

+92-310-5540387 | muhammadahsanuetm143@gmail.com | Khyber Pakhtunkhwa, Pakistan
github.com/muhammad-ahsan12 linkedin.com/in/muhammad-ahsan-6b552b279 m-ahsan.vercel.app

PROFESSIONAL SUMMARY

AI/ML Engineer with 8 months of professional industry experience building and deploying AI applications: Retrieval-Augmented Generation (RAG) pipelines, agentic workflows real-time voice agents and computer vision systems served through FastAPI and Docker. Built an end-to-end AI emergency-dispatch platform (Rescue 1122 KPK) as a final year project. Experience across education, healthcare, finance and emergency-services use cases, with hands-on work in prompt engineering, structured outputs, fine-tuning and CI/CD deployment.

TECHNICAL SKILLS

- **Generative AI & LLMs:** RAG, Agentic AI, Multi-Agent Systems, Voice AI Agents (LiveKit, WebRTC), Structured Outputs (JSON Schema), Prompt Engineering, Fine-Tuning, Transfer Learning, LangChain, LangGraph, AutoGen, Hugging Face Transformers, OpenAI API, Gemini API, Groq API
- **Machine Learning & Computer Vision:** PyTorch, TensorFlow, Keras, Scikit-learn, CNN, RNN, BiLSTM, GANs, Ensemble Models, Classification, Regression, Segmentation, Object Detection, Tracking, OpenCV, YOLO, MediaPipe, OCR (Tesseract), spaCy, NLTK
- **Vector Databases & Data:** FAISS, Pinecone, Chroma, SQLite, Pandas, NumPy, Matplotlib, Seaborn, Plotly
- **Backend & Frontend:** Python, FastAPI, asyncio, REST APIs, JWT Authentication, React, Next.js, TypeScript, Tailwind CSS, Leaflet.js, Streamlit
- **DevOps & Tools:** Docker, Docker Compose, Docker Hub, GitHub Actions (CI/CD), Git, GitHub, Google Colab, Kaggle, ClickUp, Slack
- **Professional Skills:** Problem Solving, Teamwork, Communication, Leadership, Critical Thinking, Presentation Skills, Agile Methodologies

PROFESSIONAL EXPERIENCE

Junior AI Developer | Synaptics Solution **January 2025 – Augast 2025**

- Develop and deploy chatbots, code-extraction systems and detection systems using Python, FastAPI, LangChain, TensorFlow and Transformer models.
- Lead RAG and agentic AI projects that enable context-aware, multi-step interactions and automate workflows.
- Apply fine-tuning strategies to adapt AI models to domain-specific tasks.
- Collaborate in an Agile workflow using ClickUp, Slack, Git and GitHub.

Machine Learning Engineer Intern | MakTek AI **July 2024 – September 2024**

- Built and deployed real-world chatbots using LangChain agents and Hugging Face Transformers.
- Developed agentic AI systems capable of autonomous, multi-step task execution.
- Integrated real-time inference pipelines for dynamic, context-aware user interactions.
- Applied fine-tuning and prompt engineering to improve Transformer-based model performance.

Machine Learning Intern | Mawakey **July 2024 – August 2024**

- Developed a medical diagnosis classifier using decision trees and random forests.
- Implemented a CNN image classifier and improved performance through data augmentation.
- Created a deep-learning intrusion detection system for network security using anomaly detection.

PROJECTS

Rescue 1122 KPK – AI Voice Emergency Dispatch Platform (Final Year Project) | FastAPI, LiveKit, Gemini Live, Gemini 2.5 Flash, React, TypeScript, Leaflet.js, OSRM, SQLite, Docker

- Built an end-to-end emergency response platform in which a Pashto/Urdu voice agent (LiveKit WebRTC + Gemini Live) answers the call, captures the caller's GPS location and gives first-aid guidance while help is dispatched.
- Designed an asynchronous extraction pipeline (Gemini 2.5 Flash, strict JSON schema, no-hallucination guards) that fills incident records from live transcripts every 3 exchanges, removing manual form entry.
- Implemented nearest-ambulance ranking (Haversine) with OSRM road distance, ETA and route, plus auto/manual dispatch.
- Developed a React + TypeScript + Leaflet operator dashboard and auto-generated PDF incident reports; integrated a real-time CNN + BiLSTM audio classifier.

Heart Sound Diagnostic AI (Thesis) | Python, BiLSTM, Dense layers

- Developed a machine learning project to predict potential heart conditions by analyzing heart sound recordings.
- Processed and analyzed audio data to extract relevant features for heart sound classification.
- Implemented a predictive model using Python and machine learning techniques to support preliminary cardiac assessment.

FullStack AI Assistant Chatbot | FastAPI, Next.js, TypeScript, Tailwind CSS, LangChain, LangGraph, OpenAI, Groq

- Built a full-stack AI chat application supporting text chat, image analysis, PDF Q&A and AI image generation.
- Implemented intelligent model routing, conversation memory, file uploads and a real-time interactive UI.

SalesBot – Multi-Vector RAG Chatbot | LangChain, FAISS, Groq LLaMA, Google Generative AI Embeddings, Streamlit
- Developed a customizable sales assistant that retrieves across multiple vector stores (books, FAQs, web data) for multi-source contextual answers.
- Added persona-based prompt templates, dynamic conversation memory and a branded Streamlit interface.

Multi-PDF Q&A Chatbot with CI/CD | FastAPI, LangChain, Groq LLaMA, Docker Compose, GitHub Actions, Streamlit
- Built a multi-PDF question-answering system with LangChain-based document retrieval, containerized with Docker and orchestrated with Docker Compose.
- Automated builds and deployments with GitHub Actions CI/CD workflows and a Streamlit interface for real-time Q&A.

Fraud Detection & Receipt OCR API | FastAPI, Scikit-learn, Tesseract OCR, Docker, Streamlit
- Developed a FastAPI service with a calibrated ML model for real-time credit card fraud risk scoring, integrated with Tesseract OCR for receipt extraction.
- Containerized the full pipeline with Docker and built a Streamlit interface showing fraud score, risk level and OCR results.

Domain-Specific LLM Fine-Tuning | Hugging Face Transformers, PyTorch, LangChain, Colab
- Fine-tuned LLMs on Quran translations, the Samsam dataset and customer FAQs using transfer learning, prompt engineering and parameter tuning, with custom tokenizers for dataset preparation.
- Improved accuracy and relevance for multilingual chatbots and domain-specific Q&A assistants.

Salat Pose Detection System | YOLO, MediaPipe, OpenCV
- Built a real-time computer vision tool combining YOLO object detection and MediaPipe pose estimation to detect and verify prayer postures from webcam input, with bounding-box annotation and landmark drawing.

SlideGen AI – Generative Slide Creator | LangChain, FastAPI, DALL·E
- Built an automated system that generates structured PowerPoint slides with topic-specific content and images from Pexels, Unsplash, Pollinations and DALL·E, with preview and download in the frontend.

Face & Anime Character Generation with GANs | TensorFlow, PyTorch
- Designed and trained WGAN, CGAN, StyleGAN and CycleGAN models to generate realistic human faces and stylized anime characters, using conditional generation, image-to-image translation and data augmentation.

EDUCATION

BS in Computer Science (Artificial Intelligence) **Sep 2022 – Jun 2026**

University of Engineering and Technology, Mardan | CGPA: 3.44