

MUHAMMAD SAIF UR REHMAN

+923145393465 | saif010415@gmail.com | [linkedin.com/in/muhammad-saif-ur-rehman](https://www.linkedin.com/in/muhammad-saif-ur-rehman) | github.com/Muhammad-Saif-Ur-Rehman
| muhammad-saif-ur-rehman.tech

EDUCATION

COMSATS University Islamabad Islamabad, Pakistan
Bachelor of Science in Artificial Intelligence Feb. 2022 - Jan. 2026
CGPA: 3.4/4.0

Punjab Group of Colleges, Ayub Park Campus Rawalpindi, Pakistan
Intermediate (ICS, Physics) Aug. 2019 - Aug. 2021
Marks: 992/1100

EXPERIENCE

AI/ML Engineer Dec. 2025 - Present
Freelance and Contract Projects, Self Employed Rawalpindi, Pakistan

- Built a full-stack Alzheimer's detection application: trained LeNet, VGG16, MobileNetV2, and ResNet alongside a custom model that outperformed all benchmarks and was selected as the final deliverable.
- Developed the client-facing web application using Flask (backend) and React (frontend) with real-time model inference; delivered remaining models separately for the client's thesis research.

AI Intern, Medical Imaging Lab Feb. 2025 - Jun. 2025
National Centre of AI, COMSATS University Islamabad Islamabad, Pakistan

- Fine-tuned vision LLMs on medical imaging data for classification and grounding tasks, adapting large multimodal models for domain-specific diagnostic use cases.
- Fine-tuned three open-source LLMs on custom datasets covering SQL ingestion classification, network-based query classification, and medical question-answering over clinical data.
- Independently managed end-to-end pipelines including dataset preprocessing, model configuration, training, and evaluation using Kaggle and Colab environments.

Final Year Project Lead, Fluxwell Dec. 2024 - Dec. 2025
COMSATS University Islamabad Islamabad, Pakistan

- Co-led development of a health and lifestyle advisor application as project leader in a 2-person team, architecting a system of 7 autonomous AI agents each owning a distinct health domain.
- Agents handled personalised workout and meal planning, multimodal RAG coaching, real-time vitals monitoring with anomaly alerts, live posture analysis with voice guidance, and AI-driven blog generation.
- Built a progress dashboard with visual charts and gamification features (badges, streaks) to drive long-term user engagement.

Cohort Participant Nov. 2024 - Jun. 2025
Innovista Immoquest Cohort 1 Rawalpindi, Pakistan

- Trained in Data Science, Machine Learning, Deep Learning, NLP, Computer Vision, and cloud deployment via AWS (S3, EC2, Bedrock, SageMaker).
- Built and deployed a RAG-based medical chatbot using LangChain and vector databases, gaining hands-on experience with retrieval pipelines and LLM-based response generation.

PROJECTS

Market Research Crew | *Python, CrewAI, LangChain* 2025
• Built a multi-agent CrewAI system with researcher, analyst, and writer agents that autonomously produces structured, investment-grade business validation reports in minutes, replacing hours of manual research.

AI Voice Follow-Up Agent | *n8n, Vapi, Google Sheets, Python* 2025
• Automated sales follow-ups via an n8n and Vapi workflow that reads a Google Sheet, calls unconverted leads automatically, summarizes each call with AI, and marks contacts as followed up, eliminating manual tracking overhead.

Customer Support Workflow | *n8n, Gmail, RMA Automation* 2025
• Built an n8n workflow that automates customer support end-to-end via Gmail, processes incoming queries automatically, and generates unique RMA return numbers for defective item cases.

RAG Pipeline and Medical Chatbot | *LangChain, Pinecone, FastAPI*

2024 - 2025

- Deployed a LangChain RAG chatbot with a Pinecone vector backend enabling natural-language querying of company and medical documents; built a standalone medical chatbot for clinical staff.

Pothole Detection | *YOLOv8, OpenCV, Python*

2024

- Trained a YOLOv8 segmentation model on real-world road imagery to detect and spatially segment potholes, ready for integration into automated road inspection workflows.

SQL and Medical QA Fine-Tuning | *Python, HuggingFace, Gemma LLM*

2024

- Fine-tuned Gemma LLM on labeled multi-domain SQL datasets achieving accurate classification across standard and network-based queries; separately fine-tuned on a curated medical Q&A dataset yielding improved contextual accuracy on clinical questions.

TECHNICAL SKILLS

Languages and Frameworks: Python, TensorFlow, PyTorch, Scikit-learn, FastAPI, Flask, React

Generative AI and LLMs: LangChain, LangGraph, CrewAI, OpenAI SDK, Groq API, HuggingFace, Fine-tuning, Amazon Bedrock, SageMaker

Retrieval and Vector Databases: RAG architectures, Pinecone, Qdrant, FAISS, hybrid semantic search

Computer Vision: YOLOv8, object detection, image segmentation, CNN, OpenCV, real-time inference

MLOps and Cloud: AWS (S3, EC2, SageMaker, Bedrock), Docker, DVC, MLflow, Jenkins, CircleCI, GitHub Actions

AI Agents and Automation: n8n, Vapi, multi-agent systems

Tools and Platforms: Git, GitHub, MongoDB, Streamlit, Jupyter, Google Colab, Kaggle

CERTIFICATIONS

IBM Generative AI Engineering Professional Certificate | Google Data Analytics (Coursera) | Introduction to Generative AI (Coursera) | Microsoft Office Specialist: Excel, Word, PowerPoint