

QURATUL AIN

Data Scientist | ML/AI Engineer

Email qurat-zee@hotmail.com

Portfolio [<https://quratulain20.github.io/github.io/>]

Certified Data Scientist & Machine Learning Engineer (NLP Specialist)

Driven AI practitioner with a proven record of building scalable machine learning systems and intelligent NLP-driven applications. Experienced in ML lifecycle — from model development to deployment — with a focus on solving real-world challenges using AI. Passionate about innovation, automation, and continuous learning.

EDUCATION

BSC Statistics from karachi University

INTERNSHIP

Recently completed my 25-day internship in Generative AI at Guftugu Publication, supervised by SkilledScore

Certificates

- Generative AI course with Langchain and Hugging Face (completed) from **Udemy**
- Data Science ML DL NLP boot camp from **Udemy**
- MLOPS in progress from **Udemy**

Areas of Expertise

- **Large Language Models (LLMs):** Building LLMs from scratch, fine-tuning (e.g., Llama), RAG pipelines using LangChain/LangGraph.
- **AI Agents & Automation:** Creating multi-agent systems using CrewAI and LangGraph for task automation.
- **NLP & ML Systems:** Experience with recommendation engines, anomaly detection, and text classification.

- **Cloud & MLOps:** Scalable deployment of ML models on AWS; currently exploring MLOps best practices.
 - **End-to-End AI Apps:** Developing & deploying full-stack AI solutions using Streamlit, Flask, and REST APIs.
-

Technical Skills

Languages & Libraries:

Python, PyTorch, Scikit-learn, NLTK, OpenCV

Frameworks & Tools:

LangChain, LangGraph, CrewAI, Streamlit, Flask

ML/DL Capabilities:

- Supervised & Unsupervised Learning
- Sequence Models (Seq2Seq, LSTM)
- Anomaly Detection, Recommender Systems
- Text Preprocessing Techniques, Tokenization, Embeddings

LLM & GenAI:

- RAG Pipelines | Fine-tuning LLMs (Llama)
- Embedding models (open-source & closed-source)
- Knowledge-based conversational AI
- Multi-step reasoning agents

Cloud & DevOps:

AWS (EC2, S3), Git, Docker (basic), REST APIs

MLOps (Currently learning model CI/CD, monitoring)

Projects

1. Text-to-Image Generator using Stable Diffusion

Developed a model that generates images from textual descriptions using Stable Diffusion. Utilizes advanced diffusion techniques to create high-quality visuals based on user input.

2. Multiturn Chatbot using LangChain

Created a conversational AI chatbot that can handle multi-turn dialogues.

Integrated LangChain for managing conversation context and enhancing user interactions. The chatbot is capable of understanding and responding to user queries effectively.

3. Fine-tuning GPT on Legal Documents

Fine-tuned a GPT model on legal documents to improve its understanding and generation of legal text. The model can assist with tasks such as drafting legal documents and answering legal questions.

4. AI Cover Letter Generator

Developed a tool that generates personalized cover letters based on user-provided resumes and job descriptions. Utilizes a generative AI model to create professional and compelling cover letters tailored to specific job applications.

5. Nutrition-Assistant

A conversational AI that provides dietary recommendations based on user input.

Tech Stack: LLM, LangChain, RAG, Streamlit

6. AI-Powered Hybrid Search Chatbot

Built using LangChain, this chatbot employs a hybrid search technique that combines dense and sparse embedding vectors to provide accurate answers to queries.

Tech Stack: LangChain, Groq, Streamlit, Pinecone, embedding and chat model

7. Spam and Ham Classifier

Email classifier using NLP and traditional machine learning techniques.

Tech Stack: NLTK, Scikit-learn, Pandas, Matplotlib

Currently Learning

- **MLOps:** Model versioning, CI/CD pipelines, orchestration, monitoring, deployment scalability