Tashfeen Raza

razatashfeen045@gmail.com +92 307 5286602 LinkedIn GitHub Rawalpindi, Pakistan

Summary

AI Engineer (BSAI) specializing in Agentic AI,LLM-powered automation,and RAG-based assistants using LangChain/LangGraph. Strong foundation in prompt engineering, AI evaluation, and tool-augmented LLM workflows (multi-step reasoning, function-calling, structured outputs). Experienced deploying production AI systems via FastAPI.

Education

BSc Artificial Intelligence — National University of Modern Languages (NUML), Islamabad Feb 2022 – Jan 2026
Relevant Coursework: Machine Learning, Deep Learning, NLP, Computer Vision, Speech Processing, Reinforcement Learning

Work Experience

LLM Integration & Product Enhancement Engineer — DASSOFT, NASTP Oct 2025 – Dec 2025

- Built LLM-powered agent using LangChain for security monitoring; improved incident response by 30% via prompt optimization
- Designed stress-test prompts (edge cases, multi-step reasoning) and validated outputs against expected behavior
- Deployed agent via FastAPI with structured output validation; strengthened report quality for audit use cases

AI/ML Intern — Siberkoza (Startup), NASTP Jul 2025 – Aug 2025

- Built AI-powered SOC system with Agentic AI to automate L1/L2 analyst tasks, reducing triage time by 40%
- Contributed to LLM automation prototypes using LangGraph for internal/product-facing workflows
- Supported experimentation with RAG-based assistants and retrieval-based question answering

Projects

Smart Extract - AI Surveillance System (Final Year Project) *Code available upon request*

- Designed intelligent system to detect accidents & fire from surveillance videos with severity classification (Low/High)
- Built stacked ensemble model (VGG19/InceptionV3/ResNet50) achieving 94.68% fire & 88% accident detection accuracy
- Implemented temporal voting pipeline; auto-extracts event clips with 5-sec buffers using React GUI

PhishGuard - Agentic AI Phishing Detection Chrome Extension GitHub

- Built agentic workflow with rule-based analysis, VirusTotal, AlienVault OTX, and Gemini AI for real-time phishing detection
- Developed Chrome extension with FastAPI backend; scans emails every 30s with 0-100 risk scoring using LLM reasoning
- Integrated Gmail API with OAuth 2.0; exports forensic reports to CSV with detailed threat intelligence

AI Voice Agent - Multi-Model RAG Platform GitHub — Demo

- Built enterprise voice assistant with real-time STT (Groq Whisper) and parallel LLM comparison (Gemini, Kimi, Llama 3.3)
- Implemented RAG pipeline with ChromaDB and async execution; deployed on Streamlit Cloud at $0 monthly cost

Technical Skills

LLM/Agentic Systems: LangChain, LangGraph, RAG, prompt engineering, multi-agent systems, function-calling
AI/ML Frameworks: PyTorch, TensorFlow, Hugging Face Transformers, Scikit-learn
Languages & Tools: Python, SQL, JavaScript, Git, FastAPI, Flask, Streamlit, CustomTkinter
APIs & Integrations: Gmail API, VirusTotal, AlienVault OTX, Google Gemini AI, Groq, OAuth 2.0
Databases: SQL, NoSQL, ChromaDB, FAISS

Certifications

- Machine Learning Specialization — Coursera (2024) — Generative AI with RAG & LangChain — Coursera (2024)