

MUHAMMAD ANSAB SULTAN

Senior AI Engineer

Attock, Pakistan

+92 318 0505100 • ansabsultan2002@gmail.com • linkedin.com/in/ansab-sultan • github.com/Ansab-Sultan • Portfolio

SUMMARY

Senior AI Engineer with 1.5+ years of experience, specializing in shipping 8+ production multi-agent systems in under 4 months. Expert in LangGraph, FastAPI, and Agentic RAG, delivering measurable ROI including 95% faster publishing and 90% reduced operational overhead. Proven ability to bridge complex AI research with scalable, enterprise-grade execution.

PROFESSIONAL EXPERIENCE

Senior AI Engineer, EnDevSols (Pvt.) Ltd

01/2026 – Present

Attock, Pakistan

- Owned delivery of 8 concurrent production agentic systems across distinct verticals, shipped on time with zero production rollbacks.
- Standardized LangGraph engineering patterns across a 5-person team, cutting LLM integration timelines by 30%.

Junior AI Engineer, EnDevSols (Pvt.) Ltd

07/2025 – 12/2025

Attock, Pakistan

- Designed Multi-Agent and GraphRAG architectures across 2 enterprise deployments, lifting query accuracy by 65% and automating 90% of manual knowledge-routing workflows.
- Shipped FastAPI/MongoDB microservice backends delivering 40% throughput improvement and 99.9% uptime over 6 months.

AI Engineer Intern, EnDevSols (Pvt.) Ltd

01/2025 – 06/2025

Attock, Pakistan

- Fine-tuned 3 domain-specific LLMs with hybrid RAG pipelines, improving retrieval precision by 40% and cutting hallucination rate by 60%.
- Delivered production FastAPI REST APIs with JWT and RBAC supporting 500+ concurrent users at sub-200ms p95 latency.

AI Engineer Intern, Kind AI

08/2024 – 01/2025

Attock, Pakistan

- Built end-to-end multimodal generative pipelines (text-to-image, text-to-video) powering 3 client demos in 6 months.
- Cut inference latency by 20% and GPU costs by 35% via quantization and dynamic batching with no quality regression.

PROJECTS

AI-Powered Business Incubator

SaaS

Supervisor-worker multi-agent system compressing an 8-week planning cycle into 2 hours and generating \$15K+ in business assets per run, with stateful LangGraph sub-graphs enabling zero-touch idea-to-execution validation and 99% faster time-to-market.

Tech Stack: LangGraph, LangChain, OpenAI API, FastAPI, MongoDB, Redis

AI-Powered Property Management & Triage System

PropTech

Autonomous WhatsApp triage agent that classifies unstructured tenant issues in real time, cutting call center load by 90%, backed by a 20-module management dashboard with granular RBAC eliminating 85% of manual ticket routing.

Tech Stack: LangGraph, FastAPI, MongoDB, Twilio WhatsApp API, Docker, JWT/RBAC

Production-Grade AI Publishing CMS

Multi-Agent Pipeline

Agentic multilingual editorial pipeline cutting time-to-publish by 95% across 3 languages, with role-segregated workflows and zero-touch ingestion across separate editorial and administrative pipelines.

Tech Stack: LangGraph, OpenAI API, FastAPI, MongoDB, Celery, APScheduler, Docker

Enterprise RAG Knowledge Assistant

Agentic RAG

Agentic RAG platform with autonomous tool-calling across OneDrive and Google Calendar, cutting retrieval time by 95% across 3 enterprise knowledge silos with sub-second query resolution.

Tech Stack: LangChain, Qdrant, OpenAI API, FastAPI, Neo4j, Microsoft Graph API

TECHNICAL SKILLS

Languages & Core: Python, SQL, NoSQL, Pydantic, AsyncIO, RegEx.

AI & Frameworks: LangChain, LangGraph, HuggingFace, PyTorch, OpenAI API, Ollama, vLLM, Groq, Qdrant

Backend & DevOps: FastAPI, MongoDB, Redis, Neo4j, Docker, Git, AWS, APScheduler, Celery

EDUCATION

Bachelor of Science in Artificial Intelligence

09/2021 – 07/2025

COMSATS University Islamabad

Attock Campus