

Shubhankar Poddar

Kolkata, West Bengal, India

+91-7667659289

codeaholicshub@gmail.com

LinkedIn

LeetCode

GitHub

Professional Summary

GenAI & Backend Engineer with 3 years of experience building scalable **LLM-powered applications, Multi-Agent Systems** and enterprise backend services using **Python, LangGraph, FastAPI** and **AWS**. Experienced in developing production-grade **RAG pipelines**, AI agents, and microservices for enterprise workflows. Strong problem-solving skills with **700+ DSA problems solved**, a **2100+ LeetCode Contest Rating**, and the **Guardian Badge**.

Technical Skills

Languages: Python, Java, SQL

GenAI & LLMs: LangGraph, LangChain, OpenAI API, Hugging Face, RAG, AI Agents, Prompt Engineering

Backend & APIs: FastAPI, REST APIs, Microservices, System Design

Vector Databases: Pinecone, ChromaDB

Databases: PostgreSQL, MongoDB, Redis, Elasticsearch

Cloud & DevOps: AWS, Docker, Git, GitHub Actions, CI/CD

AI Evaluation & Security: RAGAS, AI Guardrails, PII Redaction, Human-in-the-Loop Validation

Experience

Tata Consultancy Services

Aug 2023 – Jul 2026

Client: Johnson & Johnson | GenAI Engineer | Python, LangGraph, FastAPI, AWS

Kolkata, India

- **Cut compliance verification effort 60%** by architecting a production **LangGraph multi-agent system** (Researcher + Critic agents) that automated manual review workflows across a regulated healthcare enterprise.
- **Slashed LLM cost 40% and API latency from 2s to 200ms** by engineering a **Redis semantic caching** layer, directly reducing per-query infrastructure spend at enterprise scale.
- **Boosted retrieval accuracy on complex healthcare and legal corpora** by building enterprise **RAG pipelines** with **Elasticsearch hybrid search** (kNN + OpenAI embeddings) over large document sets.
- **Raised AI trust scores to 92%** by establishing an automated **RAGAS evaluation pipeline** tracking hallucination, relevance, and grounding, turning model quality into a measurable, monitored metric.
- **Shipped HIPAA-compliant AI guardrails** – PII redaction, confidence-based routing, and human-in-the-loop validation – enabling safe LLM processing of sensitive patient and legal data.
- **Scaled AI backend infrastructure** by designing asynchronous **FastAPI** microservices deployed via **Docker** on **AWS**, sustaining **99.9% uptime** while handling high-concurrency LLM inference requests without throttling.
- **Operationalized LLMOps and CI/CD pipelines** using **GitHub Actions** and integrated comprehensive observability (tracing/logging) for multi-agent workflows, accelerating production release cycles by **35%**.

Education

Madan Mohan Malaviya University of Technology

July 2019 - August 2023

B.Tech in Computer Science And Engineering(CSE)

CGPA : 7.59

Achievements

- **Winner – TCS AI Hackathon:** Secured **1st position** among **12,000+** associates globally for building an innovative GenAI solution.
- **Corporate Recognition:** Received **Special Achievement Award**, **Special Initiative Award**, and **Fresco Play Miles Award** at TCS.

Certifications

- **AWS Certified AI Practitioner**
- **AWS Certified Cloud Practitioner**