

Shivam Kumar Gupta

 +91-8092187258  shivam121799@gmail.com  LinkedIn  GitHub

Experience

Tata Consultancy Services <i>Infra Engineer – Azure, Active Directory , comm vault, TGIM</i>	July 2021 – Sep 2022 Kolkata, India
<ul style="list-style-type: none">Administered enterprise Azure AD, Intune, and Active Directory environments ensuring secure identity governance and device compliance across distributed systems.Reduced endpoint non-compliance rate by implementing automated configuration and monitoring policies; received “Contextual Master” Award for operational improvement.Collaborated with network, security, and platform teams to maintain infrastructure reliability, access control, and zero-trust aligned governance.	

Tata Consultancy Services <i>Gen AI Engineer – FastAPI, CrewAI, OpenAI API, Huggingface, RAG, MCP</i>	Oct 2022 – Present Kolkata, India
<ul style="list-style-type: none">Designed and deployed RAG pipelines combining embeddings, vector search, and LLM inference for accurate context-aware conversational systems.Built Agentic AI workflows using CrewAI and LangChain enabling autonomous task planning, execution, and multi-agent orchestration.Integrated FAISS, Pinecone, and Chroma vector databases with PostgreSQL to support scalable retrieval and structured knowledge storage.Developed and optimized FastAPI-based microservices to expose AI capabilities via REST APIs, improving response latency, throughput, and deployment reliability.	

Projects

AI Code Reviewer BOT OpenAI API , GitHub , Streamlit
<ul style="list-style-type: none">Developed an intelligent code review tool that analyzes commits, identifies potential bugs, and suggests optimizations using GPT models
RAG-Based Knowledge Chatbot Langchain , Chromadb , FastAPI , HuggingFace Embeddings
<ul style="list-style-type: none">Built a document-aware chatbot using embeddings, vector search, and prompt-chained LLM responses, improving contextual query accuracy and reducing hallucinations.
Trip Itinerary service Hugging face , Transformers, Streamlit
<ul style="list-style-type: none">Created an AI-powered travel planner that generates structured, personalized itineraries using template-driven prompts and transformer-based text generation models.

Technical Skills

Programming Languages: Python, SQL
AI & LLM Frameworks : LangChain, Hugging Face Transformers, OpenAI API, Crew AI, FastAPI, Streamlit
Vector Stores: ChromaDB, FAISS, Pinecone
Databases: PostgreSQL
Version Control/Tools: Git, GitHub, Gitlab, PgAdmin, DBeaver
Cloud: AWS

Certifications

- AWS Certified Developer Associate.
- Generative AI with LangChain & Hugging Face – Udemy

Education

Sastram Deemed to be University, Tamil Nadu
Master of Computer Application

Feb 2022 - Aug 2024
CGPA : 8.8