Tanishq Sharma

+91-7470495088 | tanishqsmail@gmail.com | github.com/tvnisxq | LinkedIn

Summary

Machine Learning Engineer specializing in **Agentic AI, LLM Orchestration, and RAG architectures**. Experienced in designing multi-agent workflows with **LangChain, LangGraph**, and integrating **LLMs**(OpenAI, Anthropic, Gemini, Groq, Mistral) into production Python backends, APIs, and microservices. Skilled in **Vector Databases**(Pinecone, ChromaDB, PGVector), prompt engineering, and scalable knowledge-retrieval system design.

Skills

Agentic AI / GenAI: LangChain, LangGraph, Multi-Agent Orchestration, Tool Integration, Model Context Protocol (MCP), Prompt Engineering, LLM Integration (OpenAI, Anthropic, Gemini, Groq)
RAG & Vector Search: Retrieval-Augmented Generation, Pinecone, ChromaDB, PGVector, Embeddings, Semantic Search
Languages & Backend: Python, Java, FastAPI, REST APIs, Microservices, Streamlit
ML/DL: TensorFlow, Keras, XGBoost, Scikit-learn, Deep Learning, Performance Optimization
Tools & DevOps: Docker, Linux, REST API Design, API Integration, Git/GitHub, GitHub Actions (CI/CD), MySQL, Postman, Pytest

Projects

LegalAdvisor | *LangChain, RAG, Pinecone, FastAPI, Groq, Llama 3.3* | GitHub Apr 2026 – May 2026

- Built a LangChain-powered RAG system for conversational querying of legal documents using Pinecone vector search and FastAPI microservices.
- Indexed 50+ legal document chunks in Pinecone and implemented semantic retrieval with LangChain, reducing irrelevant retrieval results during testing, developed scalable FastAPI microservices for ingestion, indexing, and Q&A

LangVid | *LangChain, RAG, ChromaDB, OpenAI Whisper, Sarvam Speech-to-Text* | GitHub Jun 2026 – Present

- Built a transcription-to-RAG pipeline using OpenAI Whisper, ChromaDB, and LangChain that generates summaries, action items, and grounded conversational QA from uploaded media, cutting manual review time by 70%.
- Architected a RAG-based retrieval chain with ChromaDB and LangChain for grounded follow-up Q&A on session transcripts.

ContentAlchemist | *Multi-Agent Orchestration, Langraph, Flask, Gemini-2.0, Transformers* | GitHubNov 2025 – Dec 2025

- Designed a LangGraph-based multi-agent workflow using ReAct reasoning patterns to coordinate Blog, Social, and Newsletter agents powered by Gemini 2.0.
- Reduced redundant LLM calls via a sequential orchestration pattern sharing a "Master Content Brief" context object across agents; added exponential backoff retry logic for API resilience.

Smoking Status Prediction ML Model | *Python, XGBoost, FastAPI, CI/CD* | GitHub May 2025 – July 2025

- Built an end-to-end ML system predicting smoker/non-smoker status from health indicators; engineered 20+ domain-specific features to boost predictive performance and an overall average accuracy of 74%.
- Deployed ensemble (XGBoost, Random Forest) models through FastAPI REST endpoints with automated CI/CD testing via GitHub Actions and Pytest.

Education

Lakhmi Chand Institute of Technology Bilaspur, C.G.
Bachelor of Engineering in Computer Science *2023 – 2026*

Shri Shankracharya Technical Campus Bhilai, C.G.
Bachelor of Engineering in Computer Science (AI-ML Specialization) *2022 – 2023*

Certifications

- **Introduction to Model Context Protocol** – Anthropic *Jun 2026*
- **5-Day AI Agents Intensive Course with Google** – Kaggle *Nov 2025*
- **AI Fluency Course** – Anthropic *Oct 2025*
- **8-Week Machine Learning Course** – Internshala Trainings *May 2025*

Achievements

- Selected among top 20 of 1,000+ applicants in the first round at **Caprae Capital**, a highly competitive recruitment process.
- Received 12 GitHub stars for an end-to-end Machine Learning project, reflecting community interest and recognition of the project's implementation quality.