

# Anupam Verma

Engineer 1



- 📍 Chennai, Tamil Nadu, India
- ☎ 9445864375
- ✉ anupam215769@gmail.com
- 🔗 LeetCode
- 🔗 Github
- 🔗 LinkedIn
- 🔗 Portfolio Website
- 🔗 Tableau Public

## SKILLS

### Generative AI

RAG, Fine-tuning, LangChain, LangGraph, Transformers, LLMs

### Machine Learning

TensorFlow, PyTorch, Scikit-learn

### NLP & Computer Vision

Hugging Face, NLTK, spaCy, Gensim, OpenCV

### Data Analysis & Visualization

Tableau, Power BI

### Programming

Python, C++, C

### Frontend & Backend

Chainlit, Streamlit, FastAPI, Flask

### Databases

MySQL, MongoDB

### VectorDBs / Vectorstores

ChromaDB, Qdrant, Milvus, Weaviate, FAISS

### Cloud

AWS, Azure, Docker, Git

## EDUCATION

Post Graduate Diploma in Data Science, Symbiosis Centre for Distance Learning

Jul 2023 – Jun 2025

B.Tech in CSE, Vel Tech University

Jul 2019 – Jun 2023

Passed with 9.08 CGPA

## EXPERIENCE

Comcast, Engineer 1

Jan 2023 – present

### Generative AI

- Engineered an **Agentic Retrieval-Augmented Generation (RAG)** system with **LangChain** and **LangGraph**, querying millions of code lines and cutting manual reviews by **40%** through automated suggestions, bug detection, test generation, and more efficient documentation of new RDK contributions using **Azure OpenAI gpt-4.1** and **o3 models**.
- Developed an interactive **chatbot** with **Streamlit** as the front-end and **FastAPI** as the **RESTful API** back-end, reducing latency by **60%** while creating **custom vector embeddings** through **microservices**.
- Integrated multiple agents with a **ChromaDB** vector store, combining **Similarity and Full-Text Search** to improve retrieval accuracy by **95%** and enable more precise queries of RDK documentation and codebase.
- Architected persistent **user session management** and chat history storage using **MongoDB**, enabling seamless retrieval of **past interactions** and supporting **context-aware** responses, which enhanced user satisfaction scores by **35%** and increased user engagement by **50%**.
- Integrated a **Neo4j Knowledge Graph** to map and analyze RDK **component-level dependencies**, accelerating dependency mapping by **40%** and enabling precise identification of **inter-component relationships** and creation of architecture diagrams.
- Leveraged **Selenium-based web scraping** to automate extraction of key contribution details from **Gerrit patches**, achieving **90% accuracy**, and integrated the results into **Jira via API**, increasing cross-team visibility and tracking efficiency by **70%** for over **500 code changes**.
- Deployed the end-to-end application on an **AWS EC2** instance using **Docker** containers, automating **CI/CD workflows** on **GitHub Actions** and achieving a **65% reduction** in deployment time while ensuring scalable, reliable operation with **99.9% uptime**, which facilitated smoother production.

### Tableau

- Developed **50+** interactive **RDK dashboards** in **Tableau**, including clone, code, and contribution **metrics portals**, boosting insights by **75%** and significantly improving data accessibility and efficiency for **stakeholders**.
- Leveraged advanced Tableau techniques to create visually appealing **charts, graphs, and maps**, representing **complex data** more effectively and enhancing **stakeholder** understanding by **50%**.
- Implemented innovative **data integration** with **Tableau Prep Builder** for **ETL** process like cleansing and **transformation**, **cutting data cleaning time** by **85%** and improving loading efficiency into **MySQL** while maintaining a streamlined workflow.

## PERSONAL PROJECTS

### Realtime Voice AI HR Assistant [LlamaIndex | LiveKit | LlamaParse] 🔗

- Developed a **conversational voice-based AI assistant** for HR policy queries using **gpt-4o-realtime-preview** to manage **STT-LLM-TTS** pipeline. Integrated the **LiveKit** framework for **real-time audio streaming** and **LlamaParse** for efficient **PDF parsing**, enabling accurate and natural voice interactions with **Tavus AI Avatars**.

### Customer Feedback Analysis [TensorFlow | Scikit-learn | Streamlit] 🔗

- Built a multi-class **sentiment analysis** pipeline for Amazon reviews using **NLP techniques**, comparing **Naive Bayes** (baseline 79.78%) with advanced models—**Feed-forward (Dense)**, **LSTM**, **GRU**, **BiLSTM**, **1D ConvNet**, and zero-shot **BERT**—achieving up to **88.85% accuracy**.

### SafeHome Observer [PyTorch | OpenCV | Transformers] 🔗

- Implemented a security-focused project that integrates **YOLO12 object detection** and **BLIP image captioning** to identify and describe **persons or vehicles** in real-time. By analyzing **video feeds or recordings**, it enables automated detection of **potential intrusions**, providing home surveillance.

## CERTIFICATES

- Fundamentals of Deep Learning by NVIDIA 🔗
- AWS Academy Machine Learning Foundations 🔗