Aryan Gupta

Data Science — ML — DL — NLP — AI — Agentic workflows

Ghaziabad, Uttar Pradesh

+91 80811 28481 — guptaaryan28106@gmail.com — Linkedin — GitHub

EDUCATION

Ajay Kumar Garg Engineering College   Ghaziabad, Uttar Pradesh
Bachelor of Technology in Computer Science and Engineering - 8.01   **2023 – Present (Expected 2027)**

EXPERIENCE

Big Data Center of Excellence
- Collaborated with a **15 member team**, Built and optimized ML, DL models that personalized user interactions based on behavioral patterns.
- Led the **Render 3.0 workshop**, teaching **50+** students the foundations of web development and AIML through hands-on sessions.

PROJECTS

AI Doctor at Home | *Python, LangChain, Hugging Face, FAISS, Groq LLaMA, Streamlit*   ⬡ [Visit Site]
- Engineered a **RAG system** in Python using **LangChain, LangChain-Groq**, and **Hugging Face sentence transformers**, ingested and embedded **120MB** of PDF medical textbooks,storing vectors in **FAISS**, enabling semantic search in **3 to 4** seconds response times.
- Orchestrated a **PromptTemplate** workflow with **LLaMA 3.1 (8B)**, generating precise, document-backed medical response.
- Developed and deployed an interactive **Streamlit** interface for real time user interaction,providing free home-based medical guidance..

Quora Duplicate Question Detection | *Python, RandomForest, XGBoost, Word2vec, Scikit learn, nltk, Streamlit* ⬡
- Engineered and optimized an **NLP** pipeline including **text cleaning** like **decontraction, HTML removal, stopword filtering, stemming**, and **advanced feature extraction (character distances, RapidFuzz fuzzy ratios, token similarity, Word2Vec embeddings)** to preprocess a **166MB** Quora dataset.
- Trained **RandomForest** and **XGBoost** models, **selecting XGBoost** for deployment after achieving 79.1% accuracy, delivering near state of the art performance(**industry benchmark 81%**), faster inference, and improved reliability.
- Developed and deployed a modern **Streamlit** web app integrating the full **NLP** and **ML** workflow, enabling instant duplicate question detection and enhanced user experience.

LangGraph Chatbot | *LangChain, LangGraph, LangSmith, Groq, Tools, SqliteSaver, Streamlit*   ⬡ [Visit Site]
- Engineered a **stateful** chatbot architecture using **LangGraph, Groq LLM**, and **custom-built tools (DuckDuckGo search, calculator, stock-price fetcher)**, **enabling real-time internet search, bypassing model knowledge cutoff** and supports **agentic tool** orchestration.
- Designed the **SQLite checkpointer** system with **LangGraph's SqliteSaver**, haing **persistence** conversations and **resume chat feature**, integrating **LangSmith** for **deep observability** and improved debugging.
- Developed and deployed a clean, responsive **Streamlit** UI with full streaming support, delivering instant responses and a smoother, more engaging user experience.

SKILLS

Languages: Python, Java, C, HTML, CSS
Libraries: Numpy, Pandas, Matplotlib, Seaborn, Scikit-learn, TensorFlow, Keras
Frameworks: Streamlit, LangChain, LangGraph, LangSmith
Databases: MySQL
Tools: Git, GitHub, Postman, VSCode, JupyterNotebook, Google Colab, Replit, Render, Vercel, AWS, Docker, Flask, FastAPI
Skills: Machine Learning(ML), Neural Network, Deep Learning(DL), NLP, Data Analysis, Data Science, Computer Vision, Generative AI, AI Agents

ACHIEVEMENTS & HIGHLIGHTS

- **IIT Jodhpur DevQuest Hackathon Finalist:** ML Developer Competed against **60+** teams from across India and secured a place in the **Top 13** finalists.
- **IIT Kharagpur Data Science Hackathon:** ML Developer Competed against 500+ teams nationwide and achieved a Top 30 position.
- **CodeChef Newbie** — Max Rating 1284+ [Visit Profile]