ADITYA SAGAR

162, 3rd Cross Rd, Kundalahalli Colony, Brookefield, Bengaluru, Karnataka 560037
+91-9454441867 ✉ adityasagarverma@gmail.com 🔗 Linkedin ⭘ Github

OBJECTIVE

Aspiring AI/ML Engineer with foundational skills in programming, data processing, and data generation. Experienced in multimodal data analysis and Large Language Model(LLM) applications using Retrieval-Augmented Generation (RAG) solutions, seeking to enhance data accessibility and streamline user experiences.

SKILLS

Technical Skills

- **Programming Languages** : C, Java, Python
- **Data Processing** : Data Preprocessing, Synthetic Data Generation
- **Machine Learning** : Supervised And Unsupervised Machine Learning Algorithms
- **Deep Learning** : Ann,Cnn,Rnn,NLP
- **Transformers** : Encoder-Decoder Architecture
- **Frameworks** : Langchain, Llamaindex
- **Vector Storage** : Faiss, Weaviate, Pinecone, Chromadb, Lancedb (Multimodal Vector Storage)
- **Large Language Models** : Openai, Gemini, Llama Vision,
- **Advanced Rag** : Hybrid Search/Indexing, Self-Querying, Reranking, Prompt Engineering
- **Multimodal Rag** : Video Processing Rag,Audio-Visual Rag, Multi-Document Retrieval (Pdf, Pptx With Embedded Images)

Version Control & CI/CD Fundamentals

- **Version Control** : Git, GitHub, GitHub Actions
- **Data Versioning** : DVC + Git for data versioning
- **ML Pipeline** : ML Pipeline using DVC,Experiment tracking using Mlflow,Model registration using MLflow
- **Containerization** : Docker, Docker Compose, Volumes, DockerHub
- **CI/CD Pipeline** : Continuous Integration and Continuous Deployment pipelines
- **Cloud Deployment** : Azure and AWS deployment(Containerization)

EXPERIENCE

AI/ML Engineer   **October 2023 – March 2024 (Intern), April 2024 – Present (Full-time)**
Blitzentech Solution Pvt Ltd ⬈   *Whitefield, Bangalore*

- **Intern Phase**:
  - Contributed to identifying and prioritizing feature categories to improve model accuracy by **above 80%.**
  - Conducted in-depth literature surveys to determine valuable data attributes and identified key areas for targeted data collection for above **1,000 pin code locations** in various cities.

- **Full-Time Phase**: Worked on **Ozaa** and **CT.ai** projects, focusing on the generation of synthetic datasets for over **100,000 users**.
  - Designed and implemented user profiles including details like name, address, phone number, and preferences for **10+ music genres**.
  - Delivered comprehensive datasets later utilized by the team to develop and fine-tune **content recommendation and suggestion models**.

Data Engineer Intern   **January 2022 – May 2022**
Digipodium ⬈   *Hazratganj, Lucknow*

- Acquired hands-on experience in **Deep Learning** by developing an autonomous vehicle system powered by **Raspberry Pi 4B+**.
- Achieved **70% accuracy** in steering angle prediction using a **Convolutional Neural Network (CNN)** trained on a dataset of **900 images.**
- Implemented real-time navigation using webcam image processing, integrating motor control with an **L298N driver** to enable autonomous movement.

PROJECTS

EngageAI: Automating Client Outreach ⬈ **|** LangChain, Llama 3.1, ChromaDB, Streamlit   **March 2023 – June 2023**
- Developed an automated tool using Llama 3.1 to generate personalized cold emails by analyzing job postings and extracting relevant skills and requirements.
- Extracted data from **100+ job portals**, **generated 500+ emails**, and matched **200+ portfolio links** via ChromaDB, boosting **productivity by 25%** and increasing **client response rates by 30%.**

RetailCheck AI: Product Quality Control Tool ⬈ **|** LangChain, llama-3.2-vision, ChromaDB   **August 2023 – September 2023**
- Built a tool to verify product genuineness by extracting product details **from over 100 retail websites** and comparing them with product images using large language models (LLMs).
- Leveraged Llama-3.2-90B-Vision to process product images, match them with extracted specifications, and generate analysis reports, **reducing discrepancies by 40%** and enhancing product quality control within **supply chain management (SCM).**

Analytica: Data Interpretation Tool ⬈ **|** Llama 3, Streamlit , Matplotlib, LangChain   **October 2023 – January 2024**
- A tool that Analyzes any user-provided dataset, generates contextual query responses, and **produces 6+ dynamic visualizations** (pie, bar, scatter, etc.) through an interactive chat interface with the help of Large Language Model.
- Achieved a **30% increase in data exploration efficiency**, streamlining decision-making processes for users.

Video Processing in RAG ⬈ **|** LlamaIndex, LanceDB, GPT-Vision4   **July 2023 – September 2024**
- An advanced query-answering application capable of **understanding context from over 1,000+ videos** and providing precise responses to user queries.
- Leveraging a multimodal index and LanceDB, the application streamlines video content analysis, **enhancing user interaction by 40%** and improving the overall user experience.

EDUCATION

CMR Institute of Technology	**October 2022 – October 2024**
Masters of Computer Applications - CGPA - 9.1	*Bangalore, Karnataka*
Babu Banarasi Das University	**May 2019 – August 2022**
Bachelors of Computer Applications - CGPA - 9.3	*Lucknow, Uttar Pradesh*
St Mary's School	**May 2017 – April 2019**
Intermediate - 12th - CGPA - 6.7	*Ballia, Uttar Pradesh*
St Mary's School	**May 2016 – April 2017**
SSLC - 10th - CGPA - 8.4	*Ballia, Uttar Pradesh*