Meriat Joseph

AI Engineer

meriatjoseph@gmail.com | +91 7902800574 | Github | Leetcode | Linkedin

Profile Summary

Results-driven AI/ML Engineer with hands-on experience in developing intelligent applications using Large Language Models (LLMs), Retrieval-Augmented Generation (RAG), Stable Diffusion, and LangChain. Skilled in Deep Learning, NLP, Computer Vision, and deploying scalable AI solutions with FastAPI and Streamlit. Strong in Prompt Engineering, Transformers, and MLOps practices. Adept at translating complex theory into practical innovation, committed to continuous learning and impactful problem-solving.

Experience

Nxtgen AI - AI Engineer **April 2025 - Present**

- Built and deployed end-to-end LLM-powered applications, focusing on enterprise use cases such as intelligent assistants, document search, and personalized bots.
- Implemented **Retrieval-Augmented Generation (RAG)** pipelines using **LangChain, OpenAI embeddings**, and vector databases for context-aware question answering.
- Designed and fine-tuned prompt engineering techniques to enhance the accuracy and contextual depth of generative model responses.
- Developed RESTful APIs using **FastAPI** to serve both ML and image generation models in production environments.
- Utilized **MongoDB** and **Chroma/FAISS** for storing vector embeddings and metadata for efficient search and retrieval.
- Worked on **Stable Diffusion models**, including **ControlNet** for conditional generation and **SAM-ViT-H** for high-accuracy object segmentation.
- Explored **Vision-Language Models (VLMs)** such as **Qwen-VL** for multimodal reasoning across text and images.
- Built and tested LLM workflows using open-source models, including **Meta's LLaMA 3.1 70B** for high-performance language understanding.
- Designed and deployed multi-agent RAG systems, simulating collaborative reasoning and coordination across tasks using Agentic AI frameworks.
- Deployed scalable AI workflows in a **cloud-first environment**, ensuring modularity, reliability, and integration with enterprise backends.

Projects

English Mastery – English Communication Assistance Application **Github**

An AI-powered mobile application designed to enhance English language proficiency, catering to users preparing for IELTS as well as those seeking to improve their general language skills. The app provides targeted support across the four key IELTS components — Reading, Speaking, Listening, and Writing — while also offering comprehensive lessons in grammar and vocabulary. With personalized feedback, interactive exercises, and AI-driven guidance, the app enables users to strengthen their language abilities in a structured, efficient, and user-friendly manner. This all-in one platform empowers learners to track progress, identify areas for improvement, and achieve their English language goals with confidence.

● **Comprehensive Learning**: Supports IELTS-specific modules, grammar, and vocabulary for well-rounded language development.

● **ML-Powered Intelligence**: Incorporates **LLM**, **RAG**, and **LangChain** to deliver personalized language assistance.

● **Advanced Integrations**: Utilizes **Whisper** for speech recognition, **GTTS** for text-to-speech, and **Gemini**, **llama3**,and **GPT-3.5** for AI-driven responses.

● **Embedding & Search**: Leverages **OpenAI embeddings** for efficient content retrieval and context-aware recommendations.

● **Seamless User Experience**: Built with a **Flutter** front end and **FastAPI** back end for a responsive, cross-platform experience.

● **Hosting & Deployment**: Back end hosted on **AWS**, with the app live on the **Play Store** for global accessibility.

● Integrates advanced AI support through the **Gemini API**, enhancing the application's capabilities and user interaction.

● **Technologies Used:** Flutter, FastAPI, LLM, RAG, LangChain, OpenAI Embeddings, Whisper, GTTS, Gemini, llama3, GPT-3.5, AWS, Play Store**.**

Mini Projects

● **Interactive Road Accident Analysis Dashboard:** Developed an interactive **Power BI** dashboard for road accident analysis, featuring advanced visualizations like **donut charts**, **pie charts**, **maps**, and **stacked columns**. Enabled dynamic filtering and drill-down capabilities to provide stakeholders with actionable insights and a user-friendly experience.

● **ChatBot – NLP-Powered Conversational Agent:** Github
Developed a chatbot using **NLP** techniques and **Python**, incorporating text preprocessing methods like **stemming and stopword** removal to enhance input quality. Fine-tuned a **TF-IDF** vectorization process to accurately represent textual inputs, enabling robust understanding and response generation.

● **YouTube video summarizer:** Github
Developed a YouTube video summarizer using **Generative AI** and **Streamlit**, enabling automated extraction of key points and concise summaries from video content.

Skills

- **Programming Language**: Python, R, C
- **Data Visualization**: Tableau, Power BI, Matplotlib, Seaborn
- **Machine Learning & Deep Learning**: Supervised & Unsupervised Learning, Neural Networks, LSTM, RNN, CNN,Transformers, Feature Engineering, AutoML, Model Evaluation, Data Preprocessing, Exploratory Data Analysis(EDA)
- **Generative AI & LLMs**: Large Language Models (LLMs), Generative AI, Prompt Engineering Retrieval-Augmented Generation (RAG), LangChain, Stable Diffusion, Google Generative AI, OpenAI API, Hugging Face Transformers
- **AI Agents & Agentic AI**: AI Agents, Agent Workflows, Multi-Agent Collaboration Platforms (MCP), Agentic Reasoning Systems, LangGraph
- **Natural Language Processing (NLP)**: Text Processing, Sentiment Analysis, Named Entity Recognition (NER), Language Modeling, Text Classification
- **Computer Vision**: OpenCV, Image Classification, Object Detection, Face Recognition, Diffusion Models, Image Processing, Image Segmentation

- **Frameworks & Libraries**: TensorFlow, Keras, PyTorch, Hugging Face, Scikit-learn, Pandas, NumPy
- **Web & Backend Development**: FastAPI, Streamlit, RESTful API Development, Streamlit Cloud
- **Database**s: PostgreSQL, MongoDB, Vector Embedding Storage (e.g., FAISS, Chroma)
- **Tools & Platforms**: Git, GitHub, Google Colab, VS Code, Jupyter Notebook
- **Core Competencies**: AI Deployment, Prompt Engineering, Analytical Thinking, MLOps (Basic), Problem Solving, Collaboration, Teamwork

Education

DataScience **2023-2025**

Brototype, Bengaluru

MA Econometrics **2020-2022**

Mahatma Gandhi University

CCPA – 3.96 / 5

BSc Mathematics **2017-2020**

Mahatma Gandhi University

CCPA – 7.78 /10