Thangarasu M

Data Scientist & GenAI Engineer

✉ thangamani1128@gmail.com   📞 9003787979   in linkedin   ⚆ github

PROFILE

Passionate Data Scientist and GenAI Engineer with hands-on experience in Data Analytics, Machine Learning, and Generative AI. Currently serving as a Technical Coordinator, actively mentoring students in AI and data projects. Proficient in Python, Neural Networks, Transformers, and LangChain. Continuously exploring cutting-edge technologies including LLMs, RAG pipelines, Hugging Face, and Vector Databases.

SKILLS

- Languages: Python | ML/DL: Scikit-learn, TensorFlow, Keras, PyTorch

- NLP & GenAI: NLTK,Transformers, BERT, GPT, Hugging Face, LangChain, LangGraph,RAG,Finetuing,AI Agents | RestAPIs: FastAPI, Flask, Streamlit

- Databases: MySQL, MongoDB | Vector DBs: FAISS, ChromaDB, Pinecone

- MLOps: Git, Docker, DVC, MLflow, GitHub Actions | Cloud: AWS

PROFESSIONAL EXPERIENCE

Technical Mentor – Data Science & AI   02/2025 – present
GUVI Geek Network Pvt. Ltd.

- Conduct in-person training sessions on Python programming, Machine Learning, and Generative AI technologies.
- Guide students through real-world projects involving Scikit-learn, Matplotlib, offering end-to-end support from data preprocessing to model deployment.
- Facilitate doubt-clearing sessions, interactive workshops, and personalized career mentoring for students from varied academic backgrounds.
- Actively contribute to bridging the gap between academia and industry by enhancing job readiness for over 100+ students in state-run institutions.

PROJECTS

Project: Flipkart Customer Support Bot – Headphones Category
Tech Stack: Python, LangChain, GPT-4, OpenAI Embeddings, Pinecone, FastAPI, Docker, AWS ECR, AWS App Runner, GitHub Actions
Flipkart Customer Support Bot is a domain-specific, RAG-based AI chatbot designed to assist users with detailed queries related to **headphones** listed on Flipkart. The bot reduces dependency on human support by providing instant, intelligent responses based on product data.

- Collected product-related documents for headphones and converted them into vector embeddings using **OpenAI's text-embedding-ada-002 model**.
- Stored embeddings in **Pinecone vector DB** and retrieved relevant chunks using **semantic similarity (cosine)**.
- Integrated **LangChain** to manage prompt chaining, memory handling, and RAG-based responses with **GPT-4**.
- Built the backend using **FastAPI** to serve the bot through an API layer.
- Containerized the application using **Docker** for portability and reproducibility.
- Set up **CI/CD using GitHub Actions**:
- On code push to main, Docker image is built and pushed to **Amazon ECR**
- The latest image is automatically deployed to **AWS App Runner**
- Entire system is scalable, cloud-native, and supports production-level deployment.

EDUCATION

Bachelor of Mathematics   09/2014 – 07/2017
Thiruvalluvar University

IIT-M Advanced Programming Professional & Master Data Science   09/2023 – 01/2024
IIT-M GUVI