VAMSHI KRISHNA CHIPPA

+91 8501837168 ⋄ Hyderabad, IN

chvamshikrishna8501@gmail.com ⋄ Linkedin ⋄ port-folio

OBJECTIVE

Machine Learning Engineer with 2+ years of experience in building advanced AI applications, including conversational chatbots. Skilled in leveraging Large Language Models (LLMs), prompt engineering, and deployment across different environments. Passionate about developing scalable, intelligent solutions that enhance decision-making, automate workflows, and improve user experiences. Seeking a challenging environment to apply and further deepen my expertise in AI application development and LLM optimization.

EXPERIENCE

Junior Machine Learning Engineer   **Apr 2023 – Present**
Amnet Digital, Hyderabad, India

Project: Development of a Conversational AI Visualization Chatbot with Database Interaction

- Designed and implemented an AI chatbot capable of seamless interaction with relational databases and generates SQL queries using Large Language Models (LLMs).
- Developed algorithms to process and transform data into graphical representations, enhancing data accessibility and visualization.
- Developed the application using Django framework and deployed it on AWS EC2.
- Technologies Used: Python, Django, LangChain, OpenAI, MSSQL, PostgreSQL, AWS.

Project: Development and Deployment of a Conversational AI Chatbot with Document Interaction

- Developed a conversational AI chatbot leveraging the Retrieval-Augmented Generation (RAG) framework and LLMs for seamless interaction with various document formats, including PDFs and images.
- Built and integrated backend APIs using Flask to connect the chatbot's UI and LLM models, ensuring efficient communication and data handling.
- Deployed the chatbot across multiple cloud platforms, with UI on AWS EC2 and LLM model on RunPod for optimized performance.
- Technologies Used: Python, Huggingface, SuryaOCR, Flask, AWS, RunPod, OpenAI, RAG Framework.

Project: Development of Optimization Algorithms for Sales and Inventory

- Developed optimization algorithms to enhance sales and inventory planning for the next financial year using historical data.
- Utilized the PuLP library for linear programming and optimization tasks.
- Deployed the optimized algorithms as Azure Functions to enable scalable and automated decision-making.
- Technologies Used: Python, PuLP, Azure Functions.

Project: Development of Asset IQ Application

- Designed and implemented a system to extract data from uploaded PDFs, images, and audio files, generating contextual summaries and relevant tags for efficient asset management.
- Integrated a database solution for efficient storage and implemented a search functionality to facilitate tag-based asset retrieval.
- Deployed the application on-premises on a DGX box using Triton Server.
- Technologies Used: Python, OpenAI, PyMuPDF, SQL, Triton server, DGX.

Project: Financial Document Data Extraction Application

- Developed a simple application that extracts specific data points from financial documents and saves the extracted data points into a database.
- Technologies Used: Python, Flask, PyMuPDF

SKILLS

- **Programming Languages:** Python, SQL
- **Databases:** SQL (MSSQL, PostgreSQL), FAISS (VectorDB)
- **LLMs and GenAI:** OpenAI API, LangChain, Retrieval-Augmented Generation (RAG)
- **Machine Learning:** Scikit-learn, NLP (Natural Language Processing), Optimization (PuLP), Pandas, NumPy
- **MLOps:** MLFlow, Triton Server
- **Deployment:** AWS Lambda, AWS EC2, Azure Functions, RunPod, On-premises (DGX Box)
- **Cloud Platforms:** AWS, Azure
- **Backend Development:** Flask, Fast API, Django
- **DevOps:** Git, Docker, Kubernetes, CI/CD
- **Operating Systems:** Linux, Windows

EDUCATION

Bachelor of Technology(ECE), Aurora's Scientific Technological and Research Academy (JNTUH)   2018- 2022

STRENGTHS

- Problem Solving skills
- Multi-tasking and Quick learner
- Adaptability
- Team Player
- Confidence and Positive thinking