Kartheek Yakkala

(913) 636-9562

kartheekyakkala.se@gmail.com ◇ Portfolio ◇ LinkedIn ◇ Blogs ◇ Github

Education

Master's in Computer Science
University of Central Missouri — Aug 2023 - Dec 2024
Relevant Coursework: Advance Algorithms, Artificial Intelligence, Cloud Computing — **GPA** : 3.66

Bachelor of Technology, Electronics and Communications Engineering
Jawaharlal Nehru Technological University Hyderabad — Aug 2017 - Jul 2021
Relevant Coursework: Operating systems, Computer Networks, Digital Image Processing — **GPA** : 3.77

Skills

Programming Languages : Python, Java, SQL, R, Bash
Technologies : Machine Learning, Natural Language Processing, Large Language Models(LLMs), Generative AI, Retrieval-Augmented Generation (RAG), Prompt Engineering
DevOps : Continuous Integration and Deployment, Docker, Kubernetes, Helm charts, Jenkins, Github Actions, Azure Functions, Google Cloud run, Azure File share
Frameworks and Databases : PyTorch, LangChain, React, MongoDb, MySql, ChromaDB, FAISS, Pinecone, BigQuery

Work Experience

Tata Consultancy Services(TCS)
Machine Learning Operations (MLOps) Engineer — Jul 2021- Aug 2023
Client: Albert Heijn

- Worked on building end-to-end ML model that automatically resolves 30% of incidents raised in ServiceNow, reducing their manual resolution time by over 95%. Technologies: Python, Machine Learning.
- Achieved an 95% accuracy rate in incident prediction by applying Random Forest and Support Vector Machine algorithms.
- Used GitHub as version control system for code development. Performed exploratory data analysis using Pandas and Numpy.
- Integrated ML application with MongoDB using Pymongo to manage and store 10,000+ tickets.
- Migrated on-premises monolithic ML application to Azure cloud as seperate modules using microservices and container architecture, which resulted in 20% reduction of infrastructure cost. Technologies: Docker, Kubernetes, Azure DevOps.
- Utilized API technologies such as Flask and FastAPI to facilitate communication between these containers.
- Implemented Continuous Integration and Continuous Deployment processes, which increased deployment speed and efficiency by 30% .
- Worked in agile and collaborated with cross-functional teams by using tools like Jira and Confluence to track and prioritise tasks.
- Worked on Chatbot using Rasa framework and NLP models like Spacy and BERT for the users to create, update, view tickets in ServiceNow.
- Used Azure Functions to host the components of the chatbot, such as integrating with external services, processing user inputs, and generating responses

Cognizant Technology Solutions
Programmer Analyst Trainee (Intern) — Feb 2021- Jun 2021

- Worked on development and maintenance of back-end microservices using Spring Boot

- Contributed to API documentation, system architecture, and deployment processes, fostering a collaborative environment. Participated in weekly knowledge-sharing sessions and workshops.

Projects

Proximity-Based User Recommendation System

- Developed a recommendation agent to to identify nearby users with shared interests using Large Language Models (LLMs), Tools, Retrieval-Augmented Generation (RAG), and advanced prompt engineering.
- Integrated custom tools to LLM for database interaction and geospatial analysis using VertexAI agent builder to improve the accuracy and relevance of user recommendations by querying databases and performing distance calculations. Used Cloud run functions to deploy custom tools developed

Technologies used: VertexAI, VertexAI Agent builder, LangChain, Generative AI, Large Language Models (LLMs), MongoDB, geopy, google maps api.

Chat with your docs

- Developed a context-aware application powered by Large Language Models (LLMs) and LangChain.
- Applied the Retrieval-Augmented Generation (RAG) approach to utilize a provided knowledge base as context, enabling the application to engage in informed and coherent conversations.

Technologies used: HuggingFace, Transformers, ChromaDB, LangChain, Vector DBs, Vertex AI vector search, Generative AI, Large Language Models (LLMs). Project Link

Open-source contributions

LangChain: Contributed to the development and documentation of the largest framework for developing LLM powered applications. Repository, Commits

LLamaIndex: Contributed to the documentation of Gen AI framework. Repository, Commits

PandasAI: Contributed to the development of LLM project. Repository, Commits

Achievements and activities

- Received **On The Spot** award from TCS for making architect changes to decommission few servers and migrate them to azure cloud, resulted in a 40% reduction in project budget.
- Wrote blogs on Large Language Models and Generative AI applications. (Link)
- Received Merit Certificate from Government of India for scoring a perfect GPA of 10 in 10th grade.

Certifications

- Microsoft Certified: DevOps Engineer Expert (AZ-400). Certificate
- Microsoft Certified: Azure Administrator Associate (AZ-104). Certificate
- Microsoft Certified: Azure Fundamentals (AZ-900). Certificate