RAHUL MYDUR

AI Engineer

Bengaluru, Karnataka +91 6362605010 rahulmydur@gmail.com rahulmydur Rahul Mydur

TECHNICAL SKILLS:

Programming & Frameworks
- Python
- FastAPI
- Flask
- Docker
- Git

GenAI & LLMs
- OpenAI, Claude, LLaMa etc
- Hugging Face Transformers
- LangChain
- LlamaIndex
- Prompt Engineering

Vector Databases
- FAISS
- ChromaDB
- AWS - OpenSearch
- Pinecone
- PGVector

Cloud & MLOps
- AWS (Bedrock, S3)
- GCP (Vertex AI)
- CI/CD pipelines

EDUCATION:

BE in Electronics and Communication Engineering

Dayananda Sagar Collage of Engineering | Bengaluru
2020 - 2023

Diploma in Electronics and Communication Engineering

Sandur Polytechnic | Sandur, Karnataka

LANGUAGES:

English, Kannada, Hindi, Telugu

HOBBIES:

Cooking, Sci-Fi Movies, Watching Tech News

PUBLICATIONS AND BLOGS:

- Check out the Blogs - Rahulmydur
- Listen to my Resume Podcast - Link

OBJECTIVE:

AI Engineer with nearly 2 years of experience specializing in Generative AI, LLMs, building RAG pipelines with LangChain / Llama, and deploying via FastAPI. Skilled in developing scalable AI solutions, optimizing workflows, and deploying models on cloud platforms like Google Vertex AI and AWS Bedrock. Passionate about leveraging advanced AI techniques to automate, enhance decision-making, and drive business innovation.

WORK EXPERIENCE:

Associate Software Engineer

Ascendion , Bengaluru (Parent) July, 2023 – August, 2024

- Designed and optimized ETL pipelines, migrating the data using Python, Databricks, Snowflake, and ADF, reducing data processing time by 50%.
- Developed and deployed Generative AI models using LLMs to automate data transformation, efficiently migrating data from one format to another, reducing manual effort by 40%.
- Implemented an AutoML-powered image classification model on Vertex AI, improving accuracy from 30% to 70%, reducing inference time by 60%.
- Designed and deployed a predictive analytics model for diabetes risk assessment, leveraging machine learning algorithms in Python.

AI Engineer

ZS, Bengaluru August, 2024 - Present

- Designed and deployed LLM-powered AI solutions using NLP and RAG(Retrival Augmented Generation), automating 80% to 90% of data analysis in business workflows.
- Optimized GenAI architectures, conducting gap analysis that led to a 50% improvement in system scalability and a 30% reduction in latency.
- Deployed AI/ML models on GCP Vertex AI and AWS Bedrock, automating the model deployment through MLOps pipelines by reducing deployment time.
- Enhanced an enterprise-grade Text2SQL agent with vector search (FAISS, ChromaDB, PGVector), improving query retrieval accuracy by 30%.
- Built LLM-powered AI agents using Semantic Kernel, Langraph, and prompt engineering, enabling real-time decision-making and reducing query resolution time.