

BANDI SAI SATYA RAKESH

ssrakesh.band@gmail.com ♦ [linkedin](#) ♦ [Github](#) ♦ +91 9492929825

EDUCATION

National Institute Of Technology Raipur(NIT Raipur)

2019 - 2023

BTech,Electrical Engineering

EXPERIENCE

HCL TECH

Oct 2023 -Present

Technical Lead [Gen AI Developer]

Chennai, India

- Developed an application using Generative AI (Gen AI) to automate PIN configuration test case code generation for Automated Test Equipment (ATE).
- Leveraged Gemma LLM for fine-tuning to generate testcase code, deployed the solution on Hugging Face Spaces, and conducted a thorough evaluation to ensure performance and accuracy.
- Enhanced model accuracy by fine-tuning OpenAI's GPT-4 and developed a Chroma DB vector database for Retrieval-Augmented Generation (RAG) for API Finder, achieving a 96.5 score on my test set using the Code BLEU metric.
- Extensive experience in building Generative AI applications and fine-tuning large language models (LLM) using **LoRA** and **QLoRA** techniques, hands-on experience implementing vector databases such as **ChromaDB**, **Qdrant**, and Pinecone for RAG.
- Experienced in working with Azure cloud-based services and **Azure AI services**, solid experience with version control systems and tools such as Git and GitHub.
- Demonstrated expertise in the Generative AI workflow, encompassing Data Exploration, Feature Engineering, LLM Model Selection, Fine-Tuning, Few-Shot Learning, Prompt Engineering, Deployment, Evaluation, and Prompt Flows.

PROJECTS

Gita GPT: a)Fine-tuned a large language model on the Bhagavad Gita using PEFT techniques, including LoRA and 8-bit quantization. Achieved a BLEU score of 2.

b)Experimented with the Gemini model using RAG and a FAISS vector database to perform information retrieval from a Bhagavad Gita text dataset.Gained an understanding of the differences between fine-tuning and RAG

[\[Project-Link-Finetuning\]](#) [\[Project-Link-RAG\]](#) [\[Technical Publications\]](#)

ATS Resume LLM App: Developed an application optimizing resumes for ATS using gemini api, enhancing match rates for job applications.[\[Project-Link\]](#)

Wind power forecasting using machine learning: Created precise models for wind power prediction, Demonstrated success in implementing machine learning solutions for renewable energy. ([Project Link](#))

SKILLS

- **Programing Languages:** C++,Python,Data structures and algorithms
- **Frameworks and Tools:** PyTorch,TensorFlow,LangChain,Transformers,Streamlit,Git/Github
- **Core Skills :** Machine learning,Artificial intelligence,Deep Learning,natural language processing(NLP),Generative AI
- **Generative AI Technologies:** Open-source and paid LLM models (Azure OpenAI,Google Gemini Pro,Gemma,Llama2)
- **Deployment Platforms:** Azure Functions, Hugging Face Spaces
- **Vector Databases:** ChromaDB, Pinecone
- **Cloud Storage and Database Management:** Azure Blob Storage, SQL

ACHIEVEMENTS. & RESPONSIBILITY

- Solved 500+ DSA Problems on : [GeeksForGeeks](#), [Leetcode](#)
- Secured 98.02 percentile in JEE Mains and 100+ Marks in JEE Advanced
- Courses: [Internshala ML Training](#),**Databricks** :Generative AI Fundamentals,Large Language Model: Application Through Production.