NICHENAMETLA VENKATA SAI TEJA

LinkedIn | GitHub | Hugging Face

Data Scientist

SKILLS

Programming Languages – Python, R (ETL, Machine learning, NLP, Computer Vision, LLMs)

Libraries – Pytorch, TensorFlow, Keras, OpenCV, scikit-learn, CI/CD pipeline

Data Analytics – Excel, SQL, Power BI, Matplotlib, Seaborn, NumPy, Pandas

GenAI Stack – RAG, Chatbot, MCP, Agents

Tools & frameworks – Git, Linux, AWS

Work Experience

Copart (July 2024 - Present)

Role: Jr Data Scientist

- Working on wide range of AI projects from Object detection to GenAI.
- RAG: working on internal knowledge tools to improve the operations and efficiency in company.
- Chatbot: Helping our analysts and consultants to take better, accurate, quick decisions.

EDUCATION

Hindustan Institute of Technology and Science, Chennai, TN 2020- 2024

BTech CSE, **8.8 CGPA**

MPRM Sri Viswasanthi Jr college, Vijayawada, AP 2018- 2020

Math-Physics-Chemistry, **9.82 CGPA**

RESEARCH WORK

MLTE – Multilingual Text Enhancer (June 2024) Hugging Face | GitHub

Domain: Generative AI, LLMs (LLaMA2, QLoRA, transformers) Conference paper

- Proposed MLTE, a fine-tuned LLaMA2 7B model for multilingual text-to-image prompt generation.
- Normal Text-to-Image generation models like DALL-E, Midjourney and LLMs like LLaMA2 only supports English.
- Created synthetic dataset and finetuned the LLaMA2 model to support Indian languages mainly Telugu and Hindi.
- Used PEFT method like QLoRA and Supervised finetuning with synthetic prompt dataset and achieved over **60% improvement** in image generation quality in terms of prompt and image correlation.
- Presented findings to **IEEE ICAAIC 2024** conference (DOI: 10.1109/ICAAIC60222.2024.10575122)

PROJECTS

Faculty Recruitment (May 2023) Demo | GitHub

Technologies Used: KMeans clustering, XGBoost, AWS-Elastic Beanstalk, scikit-learn, numpy, pandas

- Developed a web application that streamlines faculty recruitment for colleges. This project is to solve the challenge presented by the Ministry of Education in the 2023 Smart India Hackathon (SIH).
- Built a user-friendly interface, leveraged AWS cloud infrastructure with CI/CD pipeline.
- Reduced faculty recruitment time from **weeks to 3 days** with more than **80% accuracy** by utilizing Machine Learning techniques (KMeans Clustering, XGBoost).

CERTIFICATIONS

- Generative AI Fundamentals by Google Cloud
- Machine Learning & Deep Learning in Python & R
- Database Management Essentials by University of Colorado System