Sri Charan Thoutam

@ thoutamsricharan@gmail.com

○ CodeWithCharan

J +91 8897523345

TodeWithCharan

INTERNSHIPS

Data Science & GenAl Intern

Innomatics Research Labs

- **Sep 2024 Nov 2024**
- **Project Lead:** Led a 6-member team to develop GenAl-powered Discord bots for text, image, and video generation.
- Successfully developed two bots:
 - Boa: A conversational text generation bot leveraging LangChain and Google Gemini API, integrated with ChromaDB for personalized user sessions.
 - Kiku: An advanced bot enabling text extraction from images using PyTesseract and text-to-image generation via Hugging Face's Stable Diffusion API.
- Designed and structured a collaborative Discord server with dedicated channels for meetings, updates, GitHub logs (via webhooks), and voice discussions to streamline team communication.
- Planned future expansions of the bot to include video generation and advanced multimodal capabilities.
- LinkedIn post

EDUCATION

B.Tech in Computer Science and Engineering (AI & ML) Ganapathy Engineering College

i 09/2021 - 07/2024

CGPA: 7.01/10

Diploma in Electronics and Communication Engineering VMR Polytechnic

i 05/2018 - 07/2021

CGPA: 7.5/10

Board of Secondary Education SSC - Class X ASSISI ENGLISH MEDIUM SCHOOL

6 04/2018

CGPA: 8.3/10

SKILLS

- 1. Programming Languages: Python (Proficient)
- 2. **Data Science Expertise**: Data Cleaning, EDA, Feature Engineering, Data Preparation, Model Training and Model Evaluation
- 3. **Machine Learning**: Linear Regression, Logistic Regression, Decision Trees, Random Forest, XGBoost
- 4. **Generative AI**: LangChain, API Management (e.g., GPT, Gemini), Prompt Engineering, Fine-Tuning Models, VectorDB, RAG
- 5. **Frameworks**: TensorFlow, PyTorch, Scikit-learn, FastAPI, Flask, Streamlit.
- 6. MLOps: MLFlow, Docker, GitHub Actions, DagsHub
- 7. Databases: MySQL, MongoDB, ChromaDB
- 8. Cloud: AWS EC2, ECR, S3 bucket
- 9. Web Development: HTML, CSS, JavaScript
- 10. Version Control: Git, DVC
- 11. OS: Windows, Linux

PROJECTS

Diamond Price Prediction

Tech: Python, MLFlow, Flask, Docker, CI/CD, AWS

- Developed a machine learning model to predict diamond prices based on features like carat, cut, color, clarity.
- Used AWS S3 Bucket for dataset storage and retrieval.
- Conducted Exploratory Data Analysis (EDA), including Univariate and Bivariate analysis, to understand relationships.
- Implemented an automated MLOps pipeline with stages for Data Ingestion, Validation, Transformation, Model Training, and Evaluation.
- Set up a DAGsHub Registry for MLFlow Experiment Tracking and Model Versioning.
- Built a **Prediction Pipeline** for Model Serving using **Flask App**.
- Automated CI/CD pipeline using GitHub Actions.
- Deployed the app on AWS EC2 via ECR for production use.
- G CodeWithCharan/Diamond-Price-Prediction

Text Summarizer App

Tech: Python, NLP, Docker, CI/CD, AWS

- Developed a **Text Summarization App** that will summarize any **text**, **dialogue**, **conversation** or **article**.
- MLOps pipeline: Data Ingestion, Validation, Transformation, Model Training, and Evaluation.
- Automated CI/CD pipeline using GitHub Actions & Docker.
- Deployed the Streamlit App on AWS EC2 via ECR.
- CodeWithCharan/Text-Summarizer

ACHIEVEMENTS

- 1. Earned 5 stars in Python on HackerRank
- 2. Earned 5 stars in SQL on HackerRank

ARTICLES & PUBLICATIONS

- 1. Evolution of Language Representation Techniques: A Journey from BoW to GPT Read Article
- 2. Hacking the System Design: How Search Engines Understand and Deliver Results Read Article

CERTIFICATES

- 1. Python: HackerRank
- 2. Problem Solving: HackerRank
- 3. Machine Learning with Python: freecodecamp
- 4. Generative Al: Udemy

LINKS

- leetcode.com/u/CodeWithCharan
- hackerrank.com/CodeWithCharan
- linktr.ee/CodeWithCharan
- codewithcharan.github.io/My-Portfolio