

# Sri Charan Thoutam

📍 Hyderabad, India   ✉ thoutamsricharan@gmail.com   ☎ +91 8897523345   🌐 [github.com/CodeWithCharan](https://github.com/CodeWithCharan)  
[in linkedin.com/in/CodeWithCharan](https://www.linkedin.com/in/CodeWithCharan)   🔗 [codewithcharan.github.io/My-Portfolio](https://codewithcharan.github.io/My-Portfolio)

## Education

### Ganapathy Engineering College

Sept 2021 - July 2024

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

- GPA: 7.01/10.0
- **Coursework:** DSA, Machine Learning, Data Visualization, Database Management Systems

## Skills

**Languages:** Python (3 years), SQL (2 years), Bash (3 years)

**Technologies:** AWS(EC2, ECR, S3, Lambda), LangChain, MLFlow, Docker, TensorFlow, Scikit-learn, Pandas, NumPy, Matplotlib, Seaborn, Plotly, Power BI, FastAPI, Flask, Streamlit, Git, DVC

**Data Analysis:** Data Cleaning, Exploratory Data Analysis (EDA), Statistical Analysis, Feature Engineering

## Experience

### Data Science Intern

Remote

[Innomatics Research Labs](#) 🔗

Sep 2024 – Dec 2024

- Developed and deployed machine learning models for diamond price prediction, achieving 98% accuracy.
- Conducted Exploratory Data Analysis (EDA) to uncover trends, detect outliers, and analyze relationships within datasets.
- Implemented MLOps workflows with MLFlow and Docker for scalable model deployment.
- Led a 6-member team to develop GenAI-powered Discord bots for text and image generation.
- Authored and published two in-depth Medium articles, "Hacking the System Design" and "Evolution of Language Representation Techniques," covering advanced concepts like search engine architectures, distance metrics, language vectorization techniques, and embeddings.

## Projects

### Diamond Price Prediction (Internship Project) - [GitHub](#) 🔗

Oct 2024 - Oct 2024

- Developed a machine learning model to predict diamond prices based on features like carat, cut, color and clarity. Automated MLOps pipeline with stages for Data Ingestion, Validation, Transformation, Model Training, Evaluation and Prediction. Deployed the app on AWS EC2 for production use.

### Pharmaceutical Insight Retrieval System (Personal Project) - [GitHub](#) 🔗

Jan 2025 - Jan 2025

- Developed an advanced Retrieval-Augmented Generation (RAG) pipeline with similarity search for precise query responses. It is designed to help users gain meaningful insights from research papers and documents in the pharmaceutical domain. It uses LangChain, Google Gemini, ChromaDB and Streamlit.

## Publications & Achievements

### Evolution of Language Representation Techniques

[Read Article](#) 🔗

A Journey from BoW to GPT

Nov 5, 2024

### Hacking the System Design

[Read Article](#) 🔗

How Search Engines Understand and Deliver Results

Oct 28, 2024

### Earned 5 stars in both Python and SQL on HackerRank

[HackerRank Profile](#) 🔗

### Solved 110+ problems on LeetCode

[LeetCode Profile](#) 🔗