

Sohan Venkatesh

✉ soh.venkatesh@gmail.com  [sohan-venkatesh](#)  [sohv](#)  [Website](#)  [9611768593](#)

Education

Manipal Institute of Technology Bengaluru
Bachelor of Technology in Computer Science (Specialization in AI)

Expected August 2026
CGPA: 8.40/10

Experience

Precog-at-IIITH
Research Intern

May 2025 – Present
Hyderabad

- Studying the effect of narrow fine-tuning on LLM personas under the guidance of Prof. Ponnurangam Kumaraguru.
- Reproduced the emergent misalignment paper on Qwen2.5 models, observing misaligned behavior.

Fellowship.ai
Data Science Fellow

July 2024 – September 2024
Remote

- Developed a Jupyter copilot extension to help data scientists code, debug and test, improving efficiency by **30%**.
- Built CrewAI agents for data preprocessing and model training using GPT-4, reducing latency by **30%** and improving code readability by **50%**.

Infosys
Project Intern

May 2024 – July 2024
Hybrid

- Developed an ML model using PyCaret and Scikit-learn to predict obesity risk across multiple categories, achieving **83%** accuracy and **82%** precision.
- Utilized Pandas and Matplotlib for data analysis, optimizing model training by **30%** and built a Streamlit web app for real-time prediction.

Projects

Adaptive Prioritization in To-Do Lists | OpenAI Gymnasium, Deep Q-Network

- Developed a RL agent to dynamically prioritize tasks in a to-do list, improving user productivity.
- Implemented a A3C architecture using PyTorch to learn task prioritization based on user's behavior.

Financial Compliance Monitoring | Neo4j, Knowledge Graphs

- Developed a semantic search + rule-based prototype that monitors transaction compliance against rules stored in Neo4j graph database.
- Utilized Gemini 2.0 Flash to extract rules with extraction similarity of **80%** and ingested them into knowledge graph.

Word Embeddings from Scratch | Co-occurrence matrix, SVD

- Generated word embeddings using a co-occurrence matrix and evaluated them against pre-trained embeddings.
- Applied truncated SVD for dimensionality reduction to ensure fair evaluation and performed cross-lingual alignment of English and Hindi embeddings.

RL-based Trading Agent | PPO, Gymnasium

- Developed a PPO-based trading agent for the Sensex index, trained on 15 years of daily historical data.
- Achieved an estimated annualized return of **9%** with a Sharpe ratio of **1.2**, demonstrating effective trade execution and risk management.

SQL Chatbot | Llama, Groq, Streamlit

- Developed a SQL chatbot using Llama-3.3 on Groq, optimizing query generation by **40%** and deployed via Streamlit.
- Fine-tuned Llama-3.3 with QLoRA, increasing SQL query accuracy to **91%**, thereby optimizing model performance.

Technical Skills

Languages: Python, C, C++, JavaScript, SQL
Libraries: React, PyTorch, Tensorflow, Pandas, Numpy, Keras, Scikit-learn, Flask, OpenCV, Matplotlib, Langchain, Tesseract, PaddleOCR, Openai
Software/Tools: AWS, Azure, Git, Docker, Jupyter Notebook, MongoDB, MySQL, PostgreSQL