HARTHIK MANICHANDRA VANUMU

harthik7777@gmail.com • https://www.linkedin.com/in/harthik-mv/ • https://github.com/Harthik777

Education

Manipal Institute of Technology (MIT) Bengaluru

B.Tech in Computer Science and Engineering (Artificial Intelligence) 07/2023 - 07/2027

Key Achievements

1st Place, RoboRun Competition & TechTatva Selection 09/2024 - 10/2024

- Won MAHE Bengaluru's university-wide 'RoboRun' line follower robotics competition (1st Place).
- Subsequently, selected for the official, institutionally-funded MIT-Bengaluru team at TechTatva (MIT-Manipal's technical fest).

Volunteering

IEEE

Student Branch Executive 01/2025 - Present

- Led and managed IEEE Student Branch activities, including events and workshops, and boosted member engagement
- Served as a liaison between the IEEE Student Branch and the IEEE CS Society, improving communication and collaboration efficiency.

Radar (The Robotics Club of the College)

Outreach Head 10/2024 - Present

- Led a 5-person outreach team, securing speakers/MOUs/competition entries and organising an industry visit, and assisted the robotics team.

Skills and Competencies

Programming Languages: Python • Java • C

Data Analysis & Visualisation: Data Analytics • Statistical Analysis • Pandas • NumPy • Seaborn • Matplotlib

Machine Learning & AI:

Machine Learning Model Development & Evaluation • Feature Engineering • Data Augmentation (SMOGN) • Regression • Ensemble Learning • Scikit-Learn • LLM - Generative AI

Databases: SQLAlchemy • MySQL

Web Development & Automation: Flask • HTML • CSS • Jinja2 • Selenium • BeautifulSoup4 • Requests

Tools: Jupyter • VS Code • Git/GitHub • Excel

Projects

Predictive Modelling & Data Augmentation for Cricket Analytics

https://docs.google.com/document/d/14HNHQ-XNDwT1Xysy7s41vyIKMKrlKzlob1ni2wmspdM/edit?tab=t.0 10/2024 - 04/2025

- Developed and rigorously validated a complete machine learning pipeline for predictive modeling in niche sports analytics (WPL cricket), specifically addressing challenges of sparse datasets.
- Applied data augmentation techniques (SMOGN) and robust feature engineering to improve prediction accuracy for season-long batting performance.
- Authored a research paper detailing the methodology and findings, collaborating with faculty from the IT and Physical Education & Sports departments (MIT-Bengaluru), and received guidance from a professor at Sunway University, Malaysia.
- Evaluated diverse regression models (Gradient Boosting, XGBoost, etc.) using multi-seed cross-validation, demonstrating strong predictive performance.
- Created an automated Python CLI tool for efficient data scraping (from cricsheet.org) and preprocessing, preparing data for the modeling pipeline.

RBI NEFT Data Analysis and Visualization Platform

https://github.com/Harthik777/rbi_neft_graphs 03/2025 - 04/2025

- Developed a Flask web application allowing users to filter, query, analyze, and visualize large-scale RBI NEFT transaction datasets spanning multiple years (2016-Present) and numerous participating banks.
- Engineered an end-to-end data pipeline: automated web scraping (Selenium, BeautifulSoup) to fetch RBI data links, downloaded Excel files (Requests), processed data (Pandas), and stored structured data in a MySQL database (SQLAlchemy, PyMySQL).
- Implemented server-side data analysis (Pandas, SQLAlchemy) and generated dynamic visualizations (Matplotlib) of transaction trends (monthly volume/value) and bank rankings (top 10 by count/amount).
- Created interactive frontend views using HTML, CSS, and Jinja2 templating to display filtered data tables and generated graphs.

Publications

A Validated Machine Learning Framework for Data-Scarce Regression: SMOGN Augmentation in Cricket Performance Prediction 10/2024 - 04/2025

Status: Manuscript Draft Completed, Pending Submission (Link to draft in 'Projects' section above)