HARTHIK MANICHANDRA VANUMU

+91-7676452975 • harthik.mitblr2023@learner.manipal.edu • https://www.linkedin.com/in/harthik-mv/ • https://github.com/Harthik777

https://portfolio-website1-harthik-s-projects.vercel.app/

Education

Manipal Institute of Technology (MIT) Bengaluru
B.Tech in Computer Science and Engineering (Artificial Intelligence)   07/2023 - 07/2027

Experience

Research Intern, MIT Bengaluru   08/2024 - 05/2025

Supervisor: Dr. Usha Moorthy

- Co-authored and served as the first author on a research paper accepted at IEEE TENCON 2025, a flagship conference of IEEE Region 10.
- Engineered a joint hyperparameter optimization framework in Python that integrates data augmentation (SMOGN) and model training, demonstrating a significant improvement over traditional sequential methods.
- Trained and evaluated nine regression models, achieving a top predictive accuracy with a mean test R^2 of 0.9075 on a data-scarce sports analytics dataset.

Publications

https://drive.google.com/file/d/1E0wXnCWIiYp41ak6_RFMnBsShjpuMs3s/view?usp=sharing   10/2024 - 05/2025
A Machine Learning Framework for Data-Scarce Regression using SMOGN with Joint Hyperparameter Optimization: A Case Study with Cricket Performance Prediction

First Author on publication accepted at IEEE TENCON 2025 (flagship conference of IEEE Region 10, Asia Pacific)

https://drive.google.com/file/d/1PFSfZOt-o_5JtD59ij36Xl4SRr8nbL_6/view?usp=sharing   05/2025 - 08/2025
Performance Degradation of Deep and Traditional Machine Learning Models Under Market Volatility: An Inter-Day Stock Trend Comparison

Status: Manuscript completed and submitted for review at IEEE INDICON 2025 (the flagship IEEE India Council conference in India).

Volunteering

IEEE
Student Branch Vice-Chair   Present

- Led a student branch comprising 10 IEEE technical societies, organizing interdisciplinary events and fostering collaboration among diverse technical chapters.

Skills and Competencies

Programming Languages: Python · C · Java

Data Analytics: Statistical Analysis · Pandas · NumPy · Seaborn · Matplotlib

AI/ML: Machine Learning Model Development & Evaluation · Feature Engineering · Data Augmentation · Regression · Ensemble Learning · Scikit-Learn

Databases: SQLAlchemy · MySQL

Web Development/Automation: Flask · HTML · CSS · Jinja2 · Selenium · BeautifulSoup4 · Requests

Tools/Platforms: Jupyter · VS Code · Git/GitHub · Excel · Latex

Projects

Predictive Modelling and Data Augmentation for Cricket Analytics

- Developed and rigorously validated a complete machine learning pipeline for predictive modeling in sports analytics (WPL cricket), addressing challenges of sparse datasets.
- Applied data augmentation techniques (SMOGN) and robust feature engineering to improve prediction accuracy for season-long batting performance.
- 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

- 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 visualisations (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.

Stock Market Volatility Analysis with Machine Learning and Deep Learning Models

- Benchmarked deep learning (LSTM, DNN) and machine learning models (Gradient Boosting, ensembles) on 10 years of Indian stock data (2015–2025), comparing stable vs. volatile equities.
- Quantified volatility-driven degradation, observing >20% accuracy loss and doubled MSE in market conditions.
- Analyzed performance variation across regimes, showing reduced model reliability during events such as COVID-19.