Tejas Shanbhag

672 335 5900 | tejas.venugopal.shanbhag@gmail.com | 3rd Year Statistics

Technical Skills

Languages: Python, C/C++, SQL, Java, R
Frameworks: JUnit
Developer Tools: Git, VS Code, Jupyter Lab, R Studio
Libraries: pandas, NumPy, Matplotlib, Seaborn, scikit-learn, Tensorflow, Keras, Infer, ggplot2, Cowplot

Projects

Football (Soccer) Player Similarity Model | *Python, pandas, scikit-learn* Sep 2025 – Present

- Developed a statistical model to identify football (soccer) players with similar playing profiles using 30+ advanced performance metrics (e.g., goals, xG, passing, progression, and defensive stats).
- Engineered features with per-90 normalization and standardization to enable fair cross-player comparisons.
- Implemented a cosine similarity based retrieval system to rank players by similarity, enabling role-based scouting and recruitment insights.
- Visualized player similarity clusters to highlight stylistic overlaps and differences across positions.

Hurricane Price Gouging Linear Models | *R, Statistics, Model Selection* Jul 2025 – Aug 2025

- Collaborated with a team to analyze airline ticket pricing data from the U.S. Department of Transportation to investigate potential price gouging during major hurricanes.
- Fitted multiple linear models to pricing data and compared them using AIC to identify the best-fitting model for each hurricane.
- Detected suspicious price increases by analyzing quarter-over-quarter changes and benchmarking against the same quarter in the previous year.
- Delivered a comprehensive report summarizing findings and statistical evidence of unusual airline pricing behavior during hurricane events.

Fake Bill Detection | *R, KNN, Data Visualization, Class Imbalance Handling* Feb 2024 – Apr 2024

- Developed a KNN classification model in R to detect counterfeit bills based on their physical dimensions.
- Selected predictors by comparing distributions of real and fake bills using box plots created with `cowplot`.
- Addressed severe class imbalance through upsampling with the `themis` library, reducing model bias and improving detection reliability.
- Achieved 100% accuracy on the test set, demonstrating the model's effectiveness.

Volunteer Work

Certificate Design Aug 2022 – Dec 2022

Local Government School *Mumbai, India*

- Designed certificates for school events, ensuring visually appealing layouts for student recognition.
- Maintained a task-tracking spreadsheet to coordinate effectively with other volunteers and school staff, streamlining collaboration and delivery.

Education

University of British Columbia Vancouver, BC

Bachelor of Science in Statistics *Sep 2023 – Present*

- Relevant Coursework: Linear Models, Statistical Modelling, Statistical Inference, Probability, Intermediate Data Structures and Algorithm Design, Software Construction, Introduction to Computer Systems, Relational Databases.