

Title: Predicting Team Success in the Indian Premier League Cricket 2024 Season Using Random Forest Analysis

Author Names: Swamynathan Sanjaykumar, Subhashree Natarajan, Ponnusamy Yoga Lakshmi, Farjana Akter Bobby

Abstract: Background. Random Forest is a popular machine learning algorithm used for classification and regression tasks. The study purpose is to investigate the use of Random Forest machine learning to predict the winning chances of teams in the 2024 Indian Premier League (IPL) season.

Objectives. By analyzing comprehensive player statistics, including matches played, batting and bowling averages, as well as fielding contributions, the study aims to understand the factors that influence team success in T20 cricket and provide actionable insights for team management, betting markets, and cricket enthusiasts.

Material and methods. The study involved 10 cricket teams competing in the IPL 2024 season. Player statistics and match outcomes data from previous IPL seasons were collected and cleaned, with missing values addressed through imputation. The features were based on player statistics, including both aggregate measures and calculated metrics. A Random Forest is chosen as the machine learning model, trained using preprocessed data, with features derived from player statistics as input and match outcomes as the target variable. The dataset was split into training and validation sets, using methodologies such as cross-validation to ensure a robust model performance.

Results. The Random Forest model demonstrates strong predictive performance, with a low Mean Squared Error (MSE) of 8.2174, Root Mean Squared Error (RMSE) of 2.8666, and a high R-Squared value of 0.9173, indicating approximately 91.73% variance explained in the data. Chennai Super Kings emerge as frontrunners with a predicted performance percentage of 83.4%, while teams like Punjab Kings and Lucknow Super Giants show lower predicted performance percentages, suggesting potential areas for its improvement.

Conclusions. This study reveals the effectiveness of Random Forest machine learning in forecasting IPL match outcomes based on player statistics. It provides valuable insights into team dynamics and offers actionable recommendations for team management and cricket enthusiasts. The findings enrich our understanding of IPL match dynamics, contribute to the evolution of cricket analytics, and promote greater engagement with sport, ultimately enhancing the fan experience in the IPL.

Key words: Indian Premier League, Random Forest, machine learning, team performance prediction, cricket analytics

<https://tmfv.com.ua/journal/article/view/2619>

DOI: <https://doi.org/10.17309/tmfv.2024.2.16>