"IPLWINPREDIX"

ANALYSIS AND WINNING PREDICTION IN IPL CRICKET MATCH USING MACHINE LEARNING

Lahiru Tharanga Perera

A dissertation submitted in partial fulfilment of the requirement for
Master of Science degree in Advanced Software Engineering

School of Computing
Informatics Institute of Technology, Sri Lanka
in collaboration with
University of Westminster, UK

2023

ABSTRACT

Predicting the winner of an IPL cricket match is indeed a challenging and complex task due to the inherent unpredictability of the sport. Cricket, like many other sports, involves various factors and variables that can influence the outcome of a match, making it difficult to accurately determine the winner. Factors such as individual performances, team dynamics, weather conditions, and strategic decisions all contribute to the outcome. Despite the advancements in data analysis and statistical modeling techniques, accurately determining the match result remains complex, even for experienced analysts and fans of the game.

The existing literature on this topic offers various approaches to tackle the prediction task. However, recent research has shown that machine learning (ML) techniques can be effectively applied to enhance the accuracy of cricket match predictions. The proposed prediction model intends to utilize both traditional and novel features that are deemed significant in predicting the outcome of an IPL cricket match. The model will utilize the best fitting ML algorithms to effectively analyze and process the collected data.

In the creation of the initial models, three separate machine learning algorithms were utilized: Random Forest (RF), Naïve Bayes (NB), and Support Vector Machine (SVM). The corresponding accuracy scores for these models are 0.9747, 0.7166, and 0.7782, in that order. The selected features included the Batting Team, Bowling Team, City, remaining runs, current run rate, remaining balls, required run rate, remaining wickets, and target. The RF classifier along with Gradient Boosting (GB) yielded an accuracy score of 0.9865. These results indicate that incorporating a higher number of features along with best fitting ML algorithms can lead to an enhanced level of accuracy.

Subject Descriptors:

- Applied computing → Sport → Sport analytics
- Computing methodologies → Machine learning → Machine learning algorithms → Ensemble methods → Boosting

Key words: IPL Cricket, Match Winner Prediction, Machine Learning, Classification