Marathwada Shikshan Prasarak Mandal’s

Deogiri Institute of Engineering and Management Studies, Aurangabad

SRS Report on on

IPL SCORE PREDICTION

Submitted By

Pradeep Maku (46034) ( Btech CSE A)

Abhishek Jaju (46035) ( Btech CSE A)

Vivek Soddy (46039) ( Btech CSE A )

Dr. Babasaheb Ambedkar Technological University Lonere (M.S.)



# Department of Computer Science and Engineering

Deogiri Institute of Engineering and Management Studies,

Aurangabad

# (2021- 2022)

**Abstract**

Cricket is a popular team sport played internationally. It has tremendous spectator support and the masses show great interest in predicting the outcome of games both in their one-day international as well as the modern T-20 format. The game is governed by complex rules and scoring system. Accurate prediction of winning or losing a match faces significant challenges. Multiple parameters, including cricketing skills and performances, match venues can significantly affect the outcome of a game. These diverse parameters, along with their interdependence and variance create a non-trivial challenge to create an accurate prediction of a game. In this paper, we build a prediction system that takes in historical match data, player performance as well as the scores predicted by spectator, and predicts future match events culminating in a victory or loss. Our system predicts match outcome by analyzing pre-stored match data using simple but effective K-means clustering algorithm. We describe our system and algorithms and finally present quantitative results, demonstrating the performance of our algorithms in predicting the number of runs scored, one of the most important determinants of match outcome. Keywords: Sports prediction, K-means, analysis, clustering

INDEX

Content Page No.

1. **Introduction and description of Project**
2. **Literature Review**
3. **System Features and Requirements**
4. **UML Design**

**Introduction**

England first introduced T20 Cricket in 2003. Because of its shorter format, it became very popular. Due to its popularity of high voltage action, T20 came to India also. BCCI initiated a 20-20 cricket tournament Indian Premier League (IPL) in 2008. BCCI has been organizing the IPL T20 cricket tournament every year. The use of analytical methods in various aspects of cricket including results prediction is very important. There is a huge demand for the algorithm that best predicts the result of cricket because of its popularity and huge amount of money involved in the game. Thus the analysis of IPL results becomes more important. Prediction of outcome of a match using machine learning algorithms is an important aspect in cricket. Records of the past performance of players and other related data can be analysed to create models that predicts the winning team. This model can be created using the machine learning algorithms such as Decision Tree, Naive Bayes and K-Nearest neighbour and their results can be compared based on the Evaluation Measures as accuracy, precision, recall, sensitivity and error rate.

As of now, data analysis is need for each and every fields to examine the sets of data to extract the useful information from it and to draw conclusion and as well make decisions according to the information. Data is accessed by the computer programs developed using Machine learning to build models.The algorithm first analyses the data to create a model, specifically for understanding the patterns or trends. For creating the mining model, the model is optimized by selecting parameters and iterating. To extract actionable patterns and detailed statistics, the parameters are then fed into the dataset. This work focuses on finding the meaningful information about the IPL Teams by using the functions of R Package. R reduces the complexity of data analysis as it displays the analysis results in the form of visual representations. The dataset is loaded and a set of pre-processing is done followed by feature selection. Four machine learning algorithms Decision Tree, Naive Bayes, K-Nearest Neighbour and Random Forest are applied and the results are compared to measure the accuracy, precision, recall and sensitivity. The best of the four machine learning techniques is then applied to predict the winner and visualizes the results as graphs.

Literature Review

The work done on Data Mining of Cricket dataset describes the various data mining techniques viz Decision Tree, Naive Bayes, KNN, Random Forest applied on the IPL dataset, the model is built for predicting the results of the matches.

the selection of the best team is always required by the management for best outcome. The paper provides the optimal solution to select the best team using Data Mining Techniques rather than following the traditional method which is tedious. When we are declaring a time for the particular championship it is mandatory to select the best team and so the chance of the team to be the champion becomes easy.

In previous work the authors proposes the fuzzy clustering logic. The results of the IPL batting Statistics were grouped into various clusters and it gave efficient and effective accurate results with the Data Mining Technique – Clustering. This work has been done with the help of MATLAB. The concept of clustering is used in order to classify batting statistics of the Indian Premier League which has the fuzzy data into appropriate clusters.

Raza Ul Mustafa et al presented a study on the investigation of the feasibility of using the Twitter data to forecast the results of the match. The work has been proposed to check the machine learning techniques’ effectiveness when applied on data collected to derive insight obtained from social media networks and other real world events are predicted. The techniques used in their work are Support Vector Machine, Naive Bayes Classifier and the Linear Regression. The SVM technique holds good.

Live Cricket Score and Winning Prediction work [5] describes about the building of the model which predicts the score for the chasing team and will estimate the score of the second innings of match. The proposed work uses the concepts of Linear Regression, Naive Bayes Classifier and Reinforce Learning Algorithm and gives the idea about building a system of prediction that takes the historical data and predicts the victory or loss of the forthcoming matches.

**System Features and Requirements**