

Predicting the Outcome of ODI Cricket Matches using Decision Trees and MLP Networks

Rajiv Kumar*, Jalaz Kumar†

*Assistant Professor, †Student

Department of Computer Science and Engineering

National Institute of Technology Hamirpur, India

Email: *rajiv@nith.ac.in, †jalazkumar1208@gmail.com

Abstract—Applying Data Mining & Machine Learning in Sports Analytics is a blooming sector in the field of Computer Science. After Football, Cricket is the second most popular sports with a fan base of around 2.5 billion and mostly popular in South Asia, Australia, The Caribbeans and UK. It has tremendous spectator support and the masses show great interest in predicting the outcome of games. The result of a cricket match depends on lots of in-game and pre-game attributes. Pre-game attributes like venue, past track-records, Innings(First/Second), team strength etc. and in-game attributes like Toss, run rate, wickets in hand, strike rate etc. influence a match result predominantly. In this, we have used 2 different approaches, Decision Tree & MultiLayer Perceptron Network, to predict how these factors affect the outcome of an ODI cricket match. Based on the emerged results, We have designed CricAI: Cricket Match Outcome Prediction System. Our designed tool takes into consideration the pre-game attributes like venue(home,away,neutral),innings(first/second) & ground to & predict the outcome of given match.

Keywords: Decision Tree Classifier, MLP Classifier, Neural Networks

I. INTRODUCTION

Cricket which is the world's second most popular sport after soccer, is basically a bat and ball game played between two teams of eleven players each. Each teams comes to bat and has a single inning in which it seeks to score as many runs as possible, while the other team fields. The innings ends when the total quota of deliveries, which depends on game format have turned up, or the 10 batsman has been dismissed, whichever comes first. The prime objective is to score more runs & thus Runs are the decisive factor.

Game of Cricket is a highly unpredictable in nature. Till the very last moment, it is difficult to make accurate predictions about the game. Various natural factors affecting the game output, huge betting market and enormous media coverage have given strong incentives to model this gentlemen's game from the Machine Learning perspective.

Rules of Cricket are determined by the International Cricket Council(ICC).

There are three internationally recognized formats of Cricket matches - Test match, ODI match(One Day International) and T20 match. The main difference between these three formats is the scheduled duration of the game which directly modifies the number of deliveries each team got to play in their respective innings.

Test cricket format is the longest one and is considered as the highest standard of game. Match duration is five days in which each team get to play 2 innings each. A standard test cricket day consist of 3 sessions of 2 hours each.

One Day International i.e. ODI format is of limited overs, where each team faces 300 deliveries(50 overs). ODI match is scheduled to complete in a Day or a Day/Night combination.

T20 is the shortest internationally recognized format of this game, where each team innings consist of 20 overs. This is more of an "explosive" and more "athletic" than the other two formats.

We focus our research on One Day Internationals, the most popular format of the game. Outcome of ODI match is influenced by a large no. of factors and can be predicted like all other games. We need to find the best attributes or factors that influence the match outcome. For our study we considered the factors analyzed by and ,which are proven to have a significant impact on outcome of ODI match. The factors considered for analysis include:

- **Teams Past Performance:** This factor captures the historic outcomes of all the matches played between them.
- **Ground:** This plays a vital role as teams have great track records on grounds and carry psychological superiority over the other.
- **Innings:** This factor determines which team batted first & which batted second.
- **Home Game Advantage:** This is achieved by using Venue feature, which determines whether a particular ground is home/away/neutral for each of the playing teams.

Both of our classification models are built using these factors. To predict the outcome of ODI matches we have applied two classification techniques - Decision Trees and Multi Layer Perceptron Networks. We have conducted comparative studies among various classifiers and summarized the results in this paper.

We then built a software tool called CricAI based on emerged results, which can be used to predict the outcome of any ODI match given the concerned factors as inputs. This software of ours can be of real value to the cricketers, support staff of teams and cricket analysts in terms of analyzing the future game in advance and working towards maximizing their chances of victory.

Clustering couldn't have made any contribution to our research as we dealt with multiple independent attributes, therefore placing them in clusters after finding similarity did not seem feasible.

The rest of this paper is organised as follows. Section 2 explains the approach we have used for conducting the analysis. Section 3 presents a comparative study of the classifiers used. Section 4 presents the related work in the area and the future scope associated with this approach. Section 5 gives the conclusions.