



AutoPlay - Cricket Score Predictor

Ayush Kalla¹, Nihar Karle², Sushant Wagle³, Sandeep Utala⁴

BE Student^{1,2,3}, Assistant Professor⁴

Department of Information Technology

Vivekanand Education Society's Institute of Technology, Mumbai, India

Abstract:

Cricket is a very popular sport which is played by 12 main countries. The interest for simulating the matches and predicting their outcomes is increasing day by day and because of which it enjoys a billion dollar industry. The interest of the public is generally found to be higher in one-day international games as compared to the other two formats. One of the most common methods to determine runs scored by the team batting first is the current run-rate method. And secondly, the standard method to determine the revised target for the chasing team is the Duckworth-Lewis method. Both these methods take into consideration the most important factor i.e. the resources that a given team has in terms of over's and wickets. In this project, we are trying to build a prediction model which will accurately predict the individual and hence the team performance of a given match by taking into consideration all the other important parameters mentioned above.

Key Terms: Data mining, sports prediction, analytics, regression.

I. INTRODUCTION

The sport of cricket has a known history beginning in the late 16th century. Having originated in south-east England, it became the country's national sport in the 18th century and has developed globally in the 19th and 20th centuries. International matches have been played since 1844 and Test cricket began, retrospectively recognized, in 1877. Initially the game of cricket was played only in the Commonwealth nations, however the popularity has increased exponentially such that it is now played across continents. It has the second largest viewership by population for any sport, next only to soccer, and generates an extremely passionate following among the supporters. The game of cricket is played in three formats - Test Matches, ODIs and T20s. The first official One Day International (ODI) match was played in 1971 between Australia and England at the Melbourne Cricket Ground. While ODI cricket has developed over the past 45 years (3994 matches), the general principles have remained the same. Both sides bat once for a limited time (maximum 50 overs) with the aim in the first innings to score as many runs as possible, and in the second innings to score more than the target set in the first innings. Before the commencement of the match, the captains of both the teams meet on the ground for the toss (of a coin) to determine which team will bat first. Two batsmen and eleven fielders then enter the field and play begins when a member of the fielding team, known as the bowler, delivers (i.e., bowls) the ball from one end of the pitch towards the wicket at the other end, which is guarded by one of the batsmen, known as the striker. The commercial interest in strategic planning for ensuring victory of a particular team is huge. This has motivated thorough and methodical analysis of individual and team performance, as well as prediction of future games, across all formats of the game. Various natural factors affecting the game, enormous media coverage, and a huge betting market have given strong incentives to model the game from various perspectives.

The D/L method:-The D/L method is used to set revised targets in rain-interrupted limited-overs matches depending upon the

resources available to both the teams. These are not in direct proportion to the number of overs available to be faced, as with the average run rate method of correction. Instead they depend on how many overs are to go and how many wickets are down when the interruptions occur. The number of resources available at the interruption of the game is used as a basis to calculate the revised target. The current system only relies on the current run rate i.e number of runs scored divided by the total number of overs utilized to determine the runs that a given team will score. It does not take into consideration other important factors which are of equal, if not more, importance such as match venue(home or away), number of wickets fallen as well as the rating of a given player(player's past performance against a given team). This number of important parameters along with their interdependence creates a challenging scenario in developing an accurate model of the game.

Existing System to calculate the revised score:- The Duckworth-Lewis (D/L) method is a mathematical formulation designed to calculate the target score for the team batting second in a limited overs cricket match interrupted by weather or other circumstances. The basic principle is that each team in a limited-overs match has two resources available with which to score runs (overs to play and wickets remaining), and the target is adjusted proportionally to the change in the combination of these two resources. Whenever the overs are lost due to any unforeseen circumstances like bad weather, the target required by team 2 also is revised. This is currently done by the D/L method. However, this method also has the similar disadvantages. Hence, this method also needs to be revamped.

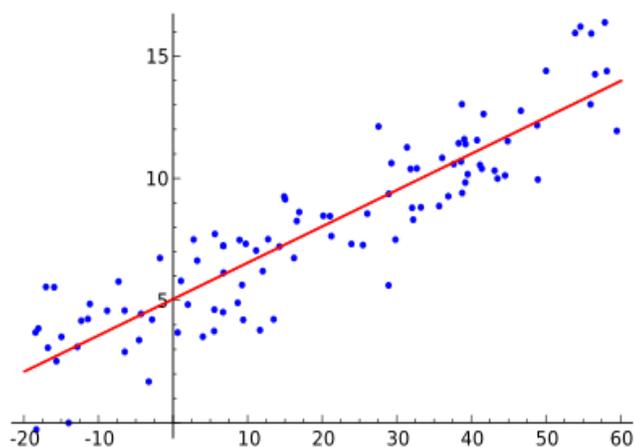
II. EXPECTED RESULTS.

At the completion of the project, cricket enthusiasts will be able to determine the approximate runs that a particular team might score on a given match day. Correspondingly, they will also be able to determine the individual performances of the

players(batsmen). Hence, the users can also determine the outcome of the match.

III. METHODOLOGY

The proposed project aims to use linear regression algorithm for predictive analysis. Linear Regression is a predictive analysis algorithm that aims to predict the future by using the previous data. It does so by plotting the past data on a graph and then uses a best fit algorithm to predict the slope and intercept of a line that is closest to most of the points. These slopes and intercepts help to predict the future outcome using the available variable in the equation of the line. Our model aims to use linear regression by taking the balls played by a player in each of his innings and the corresponding runs scored by him as the data set and passing it through the linear regression algorithm. This will give us the slope and intercept for every player and thus would enable us to predict the most probable runs a player is going to score in his next match, thus helping us to predict the total score of the team.



Algorithm

$$\text{Slope} = \frac{((\text{mean}(x) * \text{mean}(y)) - \text{mean}(x * y))}{((\text{mean}(x) * \text{mean}(x)) - \text{mean}(x * x))}$$

$$\text{Intercept} = (\text{mean}(y) - \text{mean}(x) * m)$$

Here, x is balls that a player faces in his every innings and y is the runs he scores in the corresponding innings

IV. CONCLUSION

Cricket is considered as a religion in India. Millions of fans watch cricket match in this country with many of them having tremendous interest in simulating it and more importantly in predicting the outcome of games, particularly in their one-day international format. In this project, we analyzed the work done by several people as well as the current technique of predicting the individual as well as the team performance in a one-day international cricket match. Also, we found out the shortcomings and drawbacks of each of the techniques and hence also analyzed them. The inefficiencies of the current system in place i.e. the D/L method in determining the revised targets for teams batting second and the current run-rate method to determine the amount of runs a team would score, has led to an extensive study in this field. We have tried to build our own model to determine the individual and hence the team performance of a given cricket match.

V. OBJECTIVE

Our objective in this paper is to build a prediction model that will accurately predict the runs that a team will score and correspondingly the outcome of the match. We are going to use polynomial regression to determine the value of the dependent variable i.e. the approximate number of runs that a batsman will score.

VI. REFERENCES

- [1]. Gangal, A., Talnikar, A., Dalvi, A., Zope, V. and Kulkarni, A. (2015). Analysis and Prediction of Football Statistics using Data Mining Techniques. *International Journal of Computer Applications*, 132(5), pp.8-11.
- [2]. Predicting the Outcome of ODI Cricket Matches: A Team Composition Based Approach Madan Gopal Jhanwar and Vikram Pudi *International Institution of Information Technology - Hyderabad, Gachibowli, Hyderabad, India.*
- [3]. Duckworth, FC & Lewis, AJ "Your Comprehensive Guide to The Duckworth Lewis Method for Resetting Targets in One-day Cricket", Acumen Books, 2004.
- [4]. Bower, J.M. & Beeman, D. (1995) *The Book of GENESIS: Exploring Realistic Neural Models with the General Neural Simulation System*. New York: TELOS/Springer-Verlag.
- [5]. CRICKET SCORE PREDICTION SYSTEM (CSPS) USING CLUSTERING ALGORITHM by Preeti Satao, Ashutosh Tripathi, Jayesh Vankar, Bhavesh Vaje, Vinay Varekar, *Technical Research Association India.*
- [6]. An Overview of Automated Score Recording System and Prediction System for Local Cricket Matches in India by Jayant Tikmani, Dhiren Dhameja, Hitesh Rajpal, Vivek Ahuja, Indu Dokare, *International Journal of Innovative Research in Computer and Communication Engineering (An ISO 3297: 2007 Certified Organization) Vol. 3, Issue 10, October 2015.*
- [7]. Auto-play: A Data Mining Approach to ODI Cricket Simulation and Prediction by Vignesh Veppur Sankara narayanan, Junaed Sattar and Laks V. S. Lakshmanan, *Department of Computer Science University of British Columbia Vancouver, B.C. Canada V6T 1Z4.*
- [8]. Tapas Kanungo, David M. Mount, Nathan S. Netanyahu, Christine D. Piatko, Ruth Silverman, Angela Y. Wu "An Efficient k-Means Clustering Algorithm: Analysis and Implementation" *IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE VOL. 24, NO. 7, JULY 2002*