

Intelligent Cricket Team Selection by Predicting Individual Players' Performance using Efficient Machine Learning Technique

Chetan Kapadiya, Ankit Shah, Kinjal Adhvaryu, Pratik Barot

Abstract: In any game, selection of best players in a team plays vital role in overall team performance. The team selection in any sport is the key task to ensure good performance of the team. Players are selected based on different criteria. In game of cricket selection of players should consider parameters like players own performance, ground condition, weather forecasting, opposition strength and weakness etc. Machine learning can play vital role in players' performance prediction. Machine learning uses historical data of team performance and past performance of individual players to predict overall performance of team. Prediction of individual player performance helps in team building process. Recently many researchers proposed model for prediction of player's performance for a game of cricket. Researchers' uses machine learning approach for prediction. However existing studies omits some vital features related to ground and weather in their study which have potential to make huge impact on player's performance. We performed detailed study and literature survey to propose efficient performance prediction of players for game of cricket. Our model will help in best team selection and thus improves overall team performance.

Keywords: Cricket, Performance Prediction, Machine Learning, Decision Tree, Random Forest

I. INTRODUCTION

Winning is an ultimate aim of Cricket game. Currently cricket has highest viewership after soccer [1]. Sports become more and more competitive and require through professional approach [2]. Many team introduced scientific methods to plan their strategy and decide game plan. Machine learning tool helps team management to analyze performance of rival team player. Machine learning tool helps in identification of strength and weakness of opposition and own team players [5, 11].

In recent time many researcher perform their study to propose an algorithm which can predict the performance of players [1, 2, 10, 11]. In cricket, prediction of player performance helps team management to decide final team

Revised Manuscript Received on February 26, 2020.

* Correspondence Author

Chetan Kapadiya*, PG Student, Computer Engineering Department, SVBIT, Gandhinagar, Gujarat, India. Email:cmkapadiya@gmail.com

Dr. Ankit Shah, Head of Computer Engineering / Information Technology department, SVBIT Gandhinagar, Gujarat, India. Email: shah_ankit101@yahoo.co.in

Dr. Kinjal Adhvaryu, Principal, SVBIT, Gandhinagar, Gujarat, India. Email: kinjalvk@yahoo.com

Pratik Barot, Assistant Professor, Computer Engineering Department, Government Engineering College, Gandhinagar, Gujarat, India, Email: pratikabarot@gmail.com

selection before toss.

In cricket, winning is depends upon various factors like past performances, home ground or away, experience in the match, venue specified performance, against specific team performance and the current form of the team and the form of player [3].

Today, Cricket is one of the most famous and popular sports - especially in Indian sub-continent. Many kind of natural factors affecting the game, tremendous media scope and a huge betting market have given strong incentives to model and train the game from various perspectives. As the technology is growing at a faster pace and the huge market in betting and huge demand for cricket has influenced the general population to utilize machine learning calculations to predict the results of cricket matches [6].

Use of Artificial Intelligence, machine learning, deep learning and data science makes life easier in every aspect. Use of machine learning and predicting the outcomes before the match actually played will allow the players as well as the coaches to analyze the improvement areas [7, 8]. Machine learning is booming and firmly identified with (and frequently covers with) computational insights, which also focuses on prediction-making through the use of technology. It has solid connections to numerical improvement, which hypothesis conveys strategies and application areas to the field. Machine learning is some of the time conflated with data mining where the latter subfield concentrates more on exploratory information analysis and is known as supervised learning.

Evaluating the performances of players is not a straight-forward task. Manually analysis of all the past record of each player is practically impossible. So intelligent system to predict the performance of the players based on their past record can be helpful for team management and team selectors [1, 2, 9]. Many researches have been performed to predict player performance [1, 2, 6, 10, 11]. However none of them uses weather related features which have potential to affect the performance of players. Therefore, we work on literature survey to propose an optimal methodology to predict the performances of individual players in the game of cricket match.

As most of existing researches uses small time span, our main focus is to use players past record of larger time span. Record of large time span give more accurate prediction of players' performance.

We also perform our study to predict players' performance with consideration of weather and wicket related parameters which were omitted by most of existing researchers.

II. BASIC INTRODUCTION OF CRICKET GAME

International cricket is being played in three different formats: 1. One-day International (50 overs game) , 2. T20 International (20 overs game) and 3. Test matches (5 days game) [4]. Both ODIs and T20s are considered as limited overs cricket while Test matches are of maximum 5 days. Each format consist two innings, in which each team gets one chance for bowling and one chance for batting. In ODI, one inning contains maximum of 50 overs to bowl and in T20s, as the name suggests, one inning contains maximum 20 overs to bowl. The inning gets end when the batting team has lost 10 wickets or bowling team has bowled 50 or 20 overs for ODI and T20 respectively. At the end of the first innings the teams alternating their roles, for e.g. batting team now takes the role to bowl. The Batting Team now bowls and bowling team chase a target which is set by opposite team within 50 or 20 overs for ODI and T20 respectively or before losing their 10 wickets. In the same way, the batting team now bowls and tries to stop the other team from chasing the target down within 50 or 20 overs, or by taking 10 wickets.

There are two innings in these formats, so each team has a chance to bat and bowl. In ODIs, a maximum of 50 overs can be played in one inning, and in T20s, as the name suggests, a maximum of 20 overs can be played in one inning. An inning ends when 50/20 overs have been bowled or the batting team has lost 10 wickets. At the end of the first innings, the teams change roles. The bowling team now strikes and tries to pursue the goal set by the other team within 50/20 overs or before losing its 10 goals. Similarly, the striking team now bowls and tries to prevent the other team from chasing the target within 50/20 overs or scoring 10 goals. However, test matches are played on a maximum of five days, and each team can play up to two innings in one game.

In limited cricket, the two teams have 11 players. Both teams will have to bat and bowl in one inning. Batting innings end if the batting team loses 10 wickets or 50 overs bowled by the competing team. Similarly, the next time the fielding team will strike and the other team will play. The batting team must chase the target set by the fist betting team to win the game and the opponent must restrict the batting team to chase the target to win.

Limited overs cricket is more challenging for both batsmen and bowlers [1]. Batsmen need to score runs as fast as possible and the bowlers need to restrict them by conceding the least runs and taking wickets. The focus of this thesis is the ODI format which is the most popular format in international cricket nowadays. In this study, we are trying to predict how many runs a batsman will score and how many wickets a bowler will take in a given day's match.

A. Cricket Statistics

Statistics mainly fall in two categories: Batting and Bowling. These two statistics have most influence on final outcome of match.

Batting Statistics [4]

All Statistics related to batting are:

1. Innings: The number of innings player played.
2. Not Outs: The number of times the batsman remain not out.
3. Runs: The number of runs scored by batsman.
4. Highest Score: The highest score ever made by the batsman in the past.
5. Batting Average: The batting average of batsman.
6. Centuries: The number of times batsman score 100 runs.
7. Half-Centuries: The number of times batsman score 100 runs.
8. Balls Faced: The total number of balls faced by batsman.
9. Strike Rate: $\text{Strike Rate} = (100 * \text{Runs}) / \text{Ball Faced}$
10. Run Rate: The average number of runs scores by team per 6 balls.

Bowling Statistics [4]

Statistics related to bowlers are:

1. Over: Six balls bowled by bowler is an over.
2. Maiden Overs: The number of maiden overs (In which the bowler conceded zero runs in an over) bowled.
3. Wickets: The number of wickets taken by bowler.
4. No-Balls: The number of no-balls bowled by bowler..
5. Wide: The number of wide balls bowled by bowler.
6. Bowling Average: The average number of runs conceded per wicket.
7. Strike Rate: The average number of balls bowled per wicket taken.
8. Economy Rate: The average number of runs conceded per over.
9. Five Wickets in an Innings (5w): The number of innings in which the bowler took five wickets or more.
10. Ten Wickets in a Match (10w): The number of matches in which the bowler took ten wickets.

Apart from batting and bowling there are other statistics as well which affect the final outcome of cricket match. Other important statistics are: Toss Win, Weather, Partnership, Format of the match, Venue, Pitch condition, Age of the player, Right handed or lest handed, Bowling Speed, Play in Pressure, Opponent team, Importance of match (e.g. Final Match is important than other match), One day or Day night match, Injury.

III. RELATED WORK

search provided only a few articles on the performance prediction of players in cricket [5]. Passi and Pandey [5] modeled records for hitting and bowling based on player statistics and characteristics. Other features that affect player performance, such as: the weather or the type of wickets, but could not be considered. Siripurapu [6] proposed adaptive Neuro-fuzzy model to predict players' performance but they consider only the players of a team using some batting and bowling parameters. So there is scope that rating may be improve by the use of all players at a single stretch.

Barr et al. [7], proposed criteria of comparing and selecting batsman in one day cricket match. Authors only uses batting average and strike rate for selection of a batsman. However types of wicket and weather play an important role in players' performance

In 2012, Bhattacharjee et al. [8] perform analysis to predict bowler's performance. Authors miss some of the important parameters like speed of bowling which play vital role in player's performance. This study is only to predict bowler performance.

Iyer and Sharda [9] classify bowlers and batsmen in three categories: performer, failure and moderate. Authors have followed neural network based approach for player's performance prediction. They applied this approach for the selection of World Cup 2007 team based on the individual's performance rating.

In 2017, Hasseb Ahmed et.al [10] predicted rising stars in the cricket game. They used various machine learning and mathematical approach to list top ten rising cricketers based on performance evolution and weighted average. Finally they compared rising star scores with the international cricket council rankings.

Animul et al. [11], uses SVM algorithm to predict players performance in ODI cricket match. However they use only six players for their study. SVM takes more running time when data is not linearly separable. Further it requires selection of appropriate kernel function for optimal performance.

A. Generalize Prediction Model

Many researchers proposed machine learning algorithm for players' performance prediction. Machine learning algorithms like decision tree, naïve Bayesian, SVM, random forest are used for prediction. Except random forest most of the algorithms perform sub-optimal. Generalize prediction model built from existing researches is shown in Figure-1 with the result of respective machine learning algorithm.

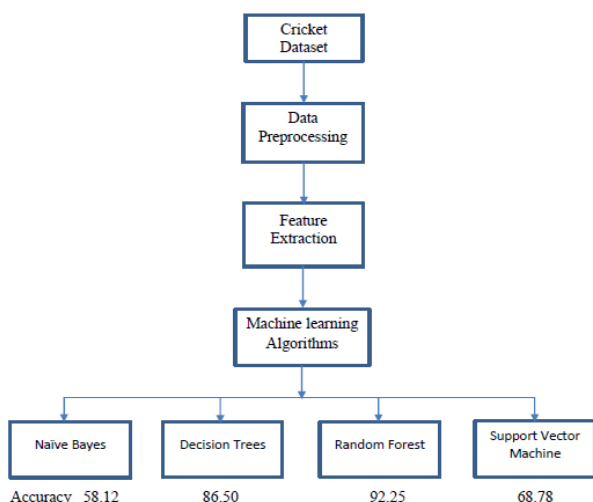


Figure 1 Generalize Prediction Model [12]

B. Problem Identification

In recent time prediction of players performance is highly desirable in competitive environment. Machine learning can play important role in this domain. Most of the existing research work omitted weather factor in

their study. Weather factor have potential to affect the performance of players and team [5]. Further most of existing study uses small set of dataset [3]. They limit their research work and concentrated on limited portion of game of cricket. Some of researchers only concentrate on batsman's performance prediction [8], while other researchers use only 4 or 5 year of data. None of the existing work have been done with consideration of holistic dataset comprise of weather, venue, wicket and cricket related statistics.

C. Proposed System Flow

Most of the existing studies only uses cricket match related statistics and omits weather related features [1, 3, 7 10]. Weather can play big role in performance of players. Flow of wind in direction of bowler running helps bowler's performance and adversely affects batsman performance. Humidity, wind flow, rain, cold in day-night match has potential to change the result of game. We proposed a system where we are using weather dataset along with cricket match statistics for player's performance prediction.

After studying literature survey on players performance prediction we found that machine learning approach with weather attribute have potential to improve prediction performance. We proposed a system in which we combine cricket dataset and weather dataset. Then we will apply feature extraction to extract most useful features. We also balance the unbalanced cricket dataset so that each class has equal influence on final prediction. Proposed system flow is given in figure-2.

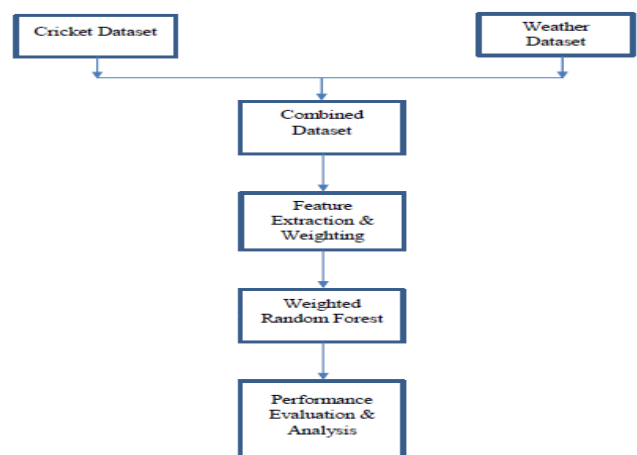


Figure 2 Proposed System Flow

IV. RESULT AND DISCUSSION

Table-1 shows result comparison of machine learning algorithms. Accuracy measure is used as performance parameter.

Table – 1, Performance (Accuracy) Comparison

Algorithm	Naïve Bayes	Decision Tree	Random Forest	SVM	Weighted Random Forest
Accuracy	58.12	86.50	92.25	68.78	93.73

As shown in table-1, in case of our proposed weighted random forest algorithm the accuracy is 93.73, which is best as compared to other traditional classification algorithms.

V. CONCLUSION AND FUTURE WORK

To design and develop an intelligent system for players' performance prediction for one day internationals cricket match it is require to consider parameters beyond the batting and bowling. Weather factor is key factor to use which have huge impact on players' performance. We proposed a machine learning model which uses cricket statistics along with weather related dataset. As combined data is imbalanced in nature we propose data balancing as preprocessing before application of machine learning algorithm. In our proposed model we had proposed model which uses new weighted random forest classifier with hyper parameter tuning and tested this model on combine balanced dataset of cricket match and weather data. We found that our proposed model had good accuracy as compare to other algorithms.

REFERENCES

1. M. G. Jhanwar and V. Pudi, "Predicting the outcome of ODI cricket matches: A team composition based approach", CEUR Workshop Proc., vol. 1842, no. September, 2016.
2. K. Passi and N. Pandey, "Increased Prediction Accuracy in the Game of Cricket Using Machine Learning", Int. J. Data Min. Knowl. Manag. Process, vol. 8, no. 2, pp. 19–36, 2018.
3. R. Lokhande and A. Professor, "Live Cricket Score and Winning Prediction", Int. J. Trend Res. Dev., vol. 5, no. 1, pp. 2394–9333.
4. [https://en.wikipedia.org/wiki/WASP_\(cricket_calculation_tool\)](https://en.wikipedia.org/wiki/WASP_(cricket_calculation_tool))
5. M. G. Jhanwar and V. Pudi, "Predicting the outcome of ODI cricket matches: A team composition based approach", CEUR Workshop Proc., vol. 1842, no. September, 2016.
6. N. Siripurapu, A. Mittal, R. P. Mukku, and R. Tiwari, "Intelligent system for team selection and decision making in the game of cricket", Smart Innov. Syst. Technol., vol. 77, pp. 467–474, 2018.
7. G. D. I. Barr and B. S. Kantor, "A criterion for comparing and selecting batsmen in limited overs cricket", J. Oper. Res. Soc., vol. 55, no. 12, pp. 1266–1274, 2004.
8. D. Bhattacharjee and D. G. Pahinkar, "Analysis of Performance of Bowlers using Combined Bowling Rate", Int. J. Sport. Sci. Eng., vol. 06, no. 03, pp. 184–192, 2012.
9. S. R. Iyer and R. Sharda, "Prediction of athletes performance using neural networks: An application in cricket team selection", Expert Syst. Appl., vol. 36, no. 3 PART 1, pp. 5510–5522, 2009.
10. H. Ahmad, A. Daud, L. Wang, H. Hong, H. Dawood, and Y. Yang, "Prediction of Rising Stars in the Game of Cricket", IEEE Access, vol. 5, pp. 4104–4124, 2017.
11. V. Q. L. N. Hdvhu et al., "Player's Performance Prediction in ODI Cricket Using Machine Learning Algorithms", 4th International Conference on Electrical Engineering and Information & Communication Technology, pp. 500-505, 2018.
12. I. P. Wickramasinghe, "Predicting the performance of batsmen in test cricket", J. Hum. Sport Exerc., vol. 9, no. 4, pp. 744–751, 2014.

AUTHORS PROFILE



Chetan Kapadiya, completed BE in Information Technology from Government Engineering College Modasa and pursuing ME in Shankersinh Vaghela Babu Institute of Technology, Gandhinagar, Gujarat, India. He has been working as an Assistant Professor in Government Engg. College, Gandhinagar, Gujarat, India.



Dr. Ankit Shah, completed Ph.D. from The M.S. University of Baroda, Vadodara, Gujarat, India. The Author's major field of study covers Big Data, Hadoop & Distributed Computing. He has been working as an Assistant Professor in Information Technology

department and Head of Computer Engineering and Information Technology Department at Shankersinh Vaghela Babu Institute of Technology, Gandhinagar, Gujarat, India.



Dr. Kinjal Adhvaryu, is working as principal of Shankersinh Vaghela Babu Institute of Technology, Gandhinagar. She has more than 15 years of experience in the field of computer science. The Author's major field of study covers Wireless Sensor Network & Ad-hoc network.



Pratik Barot, completed ME in Computer Engineering from L.D.College of Engineering Ahmedabad, Gujarat, India and pursuing Ph.D. From Gujarat Technological University, Gujarat, India. He has been working as an assistant professor in Government Engineering College, Gandhinagar, Gujarat, India. The author's major field of study covers Data Mining and Machine Learning.