A STUDY ON CRICKET TEAM SELECTION UTILIZING SOME TECHNIQUES

Suraj Jat	**Dr. Riyaz Uddin**
Research scholar	Research Guide
OPJS University	OPJS University
Department Of Physical Education	Department Of Physical Education
suraj.jat471993@gmail.com	

Abstract - *Cricket is no exception to the rule that player selection is one of the most crucial responsibilities in any sport. The rival team, the setting, the player's present form, etc. all have an impact on how well a player performs. From a roster of 15 to 20 players, the captain, coach, and team management choose 11 players for each game. To choose the best starting 11 for each game, they analyse various player traits and data. Each batter contributes by making as many runs as they can, while each bowler contribute by getting as many wickets and allowing the fewest runs possible. This essay is an effort to forecast player performance, specifically how many runs each batsman will achieve and how many wickets each bowler will claim for both sides. Both issues are intended to be classed problems, where the ranges for the amount of runs and wickets are different.*

INTRODUCTION

One of the popular methods for extracting hidden patterns from large amounts of data is data mining. The players on the squad are chosen by the sports management committee using data mining as a method to get the best outcomes. When it comes to participation, fans, and media attention, cricket is now regarded as one of the main international sports. Cricket, despite its English roots, did not garner the same level of interest and popularity throughout Europe as football did. However, it quickly gained enormous popularity in nations like India, Pakistan, Sri Lanka, Bangladesh, South Africa, Australia, New Zealand, and the West Indies, the majority of which were either British colonies or are currently under to the sway of the Crown. The International Cricket Conference (ICC) is attempting to implement new development programmes with the aim of producing more national teams capable of competing at the Test level as well as club teams that can compete in professional leagues at the national or international level in light of the growing influence and interest in the game of cricket worldwide [1].

As a result, shorter versions of the game, including the Cricket Champion League and the official Indian Premier League, were developed, including the Twenty20 World Cup in 2007. (2009).

Cricket has become a big draw nowadays due to its rising popularity and significant changes, particularly the emergence of new professional tournaments, whose performance in all of its elements is an important phenomena to observe and assess. As a consequence, more tools and software that track cricket performance are starting to appear.

Cricket is a two-team sport with eleven players on each side. Each squad has a good balance of all-rounders, ballplayers, and batters. The objective of the batsmen is to score as many runs as they can, and that of the bowlers is to take as many wickets as they can while preventing the opposing side from scoring runs at the same time. Players that are all-rounders can bat

and bowl, and they contribute by getting runs and grabbing wickets. By putting out his best effort in every game, each player helps the team function as a whole. The team he is playing against and the venue where the game is being played are two variables that affect each player's performance. It's critical to choose players who can provide their best effort in each game. A player's performance is also influenced by a number of variables, such as his present form, how well he performs against a certain team, how well he performs in a specific setting, etc. To choose the best 11 players for a particular match, team management, the coach, and the captain evaluate each player's traits, skills, and prior data.

In this study, we use supervised machine learning approaches to analyse player attributes and analytics to predict how well players would perform in matches. To do this, we forecast each batsman's and bowler's performance independently, including how many runs each will score and how many wickets each will claim throughout a certain match.

LITERATURE SURVEY

An extensive online search produced very few articles related to player's performance prediction in the game of cricket. A very small number of researchers have studied the performance of cricket players. Muthuswamy and Lam predicted the performance of Indian bowlers against seven international teams against which the Indian cricket team plays most frequently. They used back propagation network and radial basis network function to predict how many runs a bowler is likely to concede and how many wickets a bowler is likely to take in a given ODI match [2]. Wickramasinghe predicted the performance of batsmen in a test series using a hierarchical linear model. Using neural networks study of predicting "How many wickets will a bowler take?" can be possible, but their work was limited to only eight Indian bowlers and difficult to generalize for all the bowlers in the world. So some supervised machine learning algorithms to build prediction models that can be used to predict the performance of any player in a given match [3].

Iyer and Sharda[4] used neural networks to predict the performance of players where they classify batsmen and bowlers separately in three categories – performer, moderate and failure. Based on the number of times a player has received different ratings, they recommend if the player should be included in the team to play World Cup 2007. Jhanwar and Paudi[5] predict the outcome of a cricket match by comparing the strengths of the two teams. For this, they measured the performances of individual players of each team. They developed algorithms to model the performances of batsmen and bowlers where they determine the potential of a player by examining his career performance and then his recent performances. Lemmer[6] defined a new measure called Combined Bowling Rate to measure the performance of bowlers.

A. Batting And Balling Performance

Better selection tools must be created in order to choose the team's ideal lineup of players. In their study [7], Kusiak, Kern, Kernstine, and Tseng used K Nearest Neighbor to determine a score that was closer to a player's threshold or standard score. By discussing certain ODI cricket affecting variables including physical strength, scoring, home field advantages, day/night scenario, toss, and selection of batting first or second, Mickey Arthur has encapsulated the force of involvement or responsibility in a game triumph. Understanding and estimating the values utilised in Bayesian classifiers can help identify how these

variables will impact player performance in a cricket match [8]. Holder and Nevill's logistic and standard regression models were recommended to determine player rankings and the benefits of playing at home. Their study revealed that players perform better at home. Additionally, reverse data mining technology is used to identify the weak spots of other teams, allowing for the planning of defensive strategy for the players' next game [9]. The average score, strike rate, and consistency of performance are used to determine the form and selection of cricket players. Here, the top batsmen are classified in accordance with a classification system [10].

The bowling average, strike rate, and economy are three conventional bowling metrics that make up the combined bowling rate. This combined bowling rate was used by Bhattacharjee and Pahinkar [11] to evaluate the effectiveness of the bowlers in the Indian Premier League (IPL). They also identified other characteristics that have an impact on bowlers' performance and used a multivariate regression model to pinpoint the elements that are empirically accountable for bowlers' performance. Mukharjee [12] used social network analysis to evaluate the performance of a team's pitchers and batters. Using player-vs-player data accessible for test and ODI cricket, he created a directed and weighted network of batsmen-bowlers. Using the history of cricket's dismissals of batters, he also created a network of batsmen and bowlers. Shah [13] also revealed brand-new metrics for gauging player effectiveness. The new measurement for bowlers also takes into consideration the calibre of each batter they are bowling to, as well as the calibre of each bowler they are facing.

The following conclusions are drawn from a thorough assessment of the literature on game and player performance. Stretch (2003) conducted study on cricketer injuries that can impair their playing performances. In this research, the author used a hierarchical linear model (HLM) to simulate the types of injuries sustained by South African cricketers, including medical staff members and physical therapists [14]. A strategy based on a nonparametric approach was put out by Kimber and Hansford (1993) to evaluate the batting abilities of cricket batters. In all facets of the game, cricketers are known for their unfathomable capacity for reaction time [15]. Attempts were made by Balasaheb, Maman, and Sandhu (2008) to determine how visual skills training could improve a batsman's performance. In their study, they demonstrated how visual abilities enhance cricketers' response times, depth perception, and eyes, which ultimately enhances the player's batting performance [16].

Lemmer (2008) looked at player performances when players take part in a limited number of cricket matches while addressing the required modified methods to examine player performance in the game of cricket.

SYSTEM DEVELOPMENT

    A.   **Introduction to Proposed System**
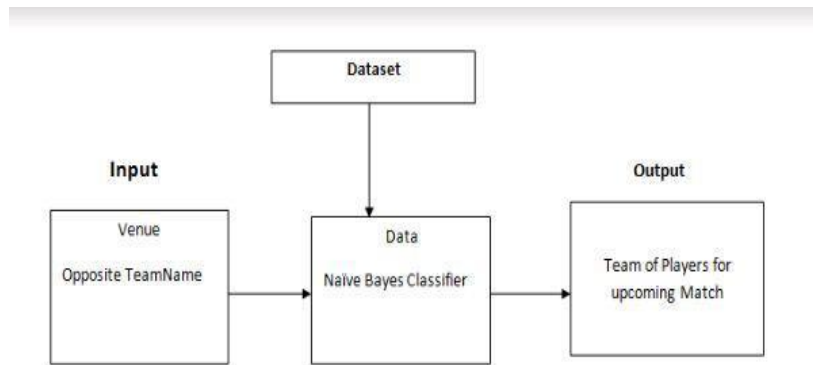
Fig 1: Proposed System

1) By examining their prior performance, the system focuses on forecasting future performance. The opposing team name and location are the model's inputs. To create the pertinent characteristics and get the choice from the model, a decision tree and induction machine learning method are utilised. The proposed system's data mining model is shown in the above figure.

2) **DATA PREPROCESSING**
Calculating the Traditional Attributes

Since player statistics like average, strike rate, and other metrics are not immediately accessible for every game, we derived them from the innings-by-innings list using mathematical formulas and aggregate functions. These characteristics are often used to gauge a player's effectiveness. The following are these qualities:

Batting Attributes

The amount of innings the hitter has played in up to game day. This characteristic denotes the batsman's level of experience. The experience level of a batter increases with the number of innings played.

No of Innings: the amount of innings the hitter has played in up to game day. This characteristic denotes the batsman's level of experience. The experience level of a batter increases with the number of innings played.

Batting Average:
The average amount of runs scored every inning is known as the batting average, or simply average. This trait reveals the player's ability to score runs.

Average = Runs Scored / Number of times dismissed
Strike Rate (SR):

The average number of runs scored per 100 balls faced is known as the strike rate. In limited overs cricket, it's critical to accumulate runs quickly. Given that the team only has a certain amount of overs, more runs scored slowly are really fairly detrimental to the team. This characteristic reveals the batsman's rate of run production.

Strike Rate: (Runs Scored / Balls Faced)
* 100
Centuries:
Instances in which a batter scores at least 100 runs. The player's ability to play more innings and score more runs is shown by this trait.

(a) **Fifties:**

Number of innings when the batter scored at least 50 runs but not more than 100. The player's ability to play more innings and score more runs is shown by this trait.

(b) **Zeros:**

No runs were scored by the hitter in the specified number of innings. This information reveals how often the batsman failed to score runs; as a result, it is a negative aspect that adversely affects the batter's projection.

(c) **Highest Score:**

The most runs a batter has ever scored in one (single) inning throughout his career. The venue attribute calculation algorithm makes use of this attribute. This characteristic demonstrates the batsman's ability to score runs at the location. A player is more likely to score more runs in a location where he has previously posted a very high score.

Bowling Attributes

(a) **No of Innings:**

(b) The number of innings in which the bowler bowled at least one ball. It represents the bowling experience of a player. The more innings the player has played, the more experienced the player is.

(c) **Overs:**

(d) the quantity of overs a bowler has bowled. This characteristic also reveals the bowler's level of expertise. The experience of a bowler increases with the number of overs bowled.

(e) **Bowling Average:**

The bowler's "bowling average" is the amount of runs allowed per wicket taken. This quality shows the bowler's capacity to prevent the batsman from both scoring runs and getting wickets at the same time. Lower bowling average numbers indicate more skill.

(f) **Bowling Average:**

Number of runs conceded / Number of wickets taken.

(g) **Bowling Strike Rate:**

The number of balls bowled per wicket taken is known as the bowling strike rate. This characteristic reveals the bowler's ability to take a wicket. Lower values indicate that the bowler can quickly take down opponents.

(h) **Strike Rate:**

Number of balls bowled / Number of wickets taken.

(i) **Four/Five Wicket Haul:**

When a bowler has more than four wickets than an innings. This quality shows the bowler's potential to record more wickets throughout an innings. The player is more competent the higher the value.

APPLICATIONS

1. Player selection for the cricket team in the IPL, where one team must play multiple teams at several venues.

2. Player selection in international competitions where one team must face many opponents, such as the ICC World Cup or ICC Champions Trophy

CONCLUSIONS

Team management finds it very challenging to choose a team, particularly in situations when one team will face many teams under diverse circumstances. Player performance varies depending on the opponent team and the location. So it takes a lot of effort to manually keep track of everything and then create a border to choose participants. The suggested technique will thus be very helpful in this regard since it will provide a comprehensive player analysis and a forecast of his performance with only one click. As a result, it will improve the precision and accuracy of the team selecting process.

FUTURE SCOPE

1. Currently, we are utilising the IPL match dataset to demonstrate the effectiveness of the proposed technique; however, in the future, we may construct other datasets to forecast players at the international level.

2. The system may be expanded to handle dynamic match data, allowing the administrator to upload the data after each match to make the analysis more fluid.

REFERENCES

[1]. R. P. Schumaker, O. K. Solieman and H. Chen,"Predictive Modeling for Sports and Gaming," in Sports Data Mining, vol. 26, Boston, Massachusetts: Springer, 2010.

[2]. S. Muthuswamy and S. S. Lam, "Bowler Performance Prediction for One-day International Cricket Using Neural Networks,"in Industrial Engineering Research Conference, 2008.

[3]. I. P. Wickramasinghe, "Predicting the performance of batsmen in test cricket," Journal of Human Sport Excercise, vol. 9, no.4, pp. 744-751, May 2014..

[4]. S. R. Iyer and R. Sharda, "Prediction of athletes performance using neural net- works:An application in cricket team selection," Expert Systems with Applica- tions, vol. 36, pp. 5510-5522, April 2009.

[5]. M. G. Jhanwar and V. Pudi, "Predicting the Outcome of ODI Cricket Matches: A Team Composition Based Approach," in European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (ECMLP- KDD 20162016), 2016.

[6]. H. H. Lemmer, "The combined bowling rate as a measure of bowling performance in cricket,"South African Journal for Research in Sport, Physical Education and Recreation, vol. 24, no. 2, pp. 37-44, January 2002..

[7]. A. Kusiak, J. A. Kern, K. H. Kernstine, and B.T. L. Tseng, "Autonomous decision-making: A data mining approach," IEEE Transactions on Information Technology in Biomedicine, vol. 4, pp. 274–284, 2000.

[8]. "www.smh.com.au," Pietersen tells of racial discrimination-Cricket-Sport. [Online]. Available: http://www.smh.com.au/n ews/cricket/pietersen-tellsof racialdiscrimination/2006/08/30/1156816967 716.html. [Accessed: 10- Mar- 2018].

[9]. R. L. Holder and A. M. Nevill, "Modelling performance at international tennis and golf tournaments: is there a home advantage?" Journal of the Royal Statistical Society:Series D (The Statistician), vol. 46, no. 4, pp. 551–559, 1997.

[10]. H. H. Lemmer, "A measure for the batting performance of cricket players : research article," South African Journal for Researchin Sport, Physical Education and Recreation, vol. 26, no. 1, 2004

[11].D. Bhattacharjee and D. G. Pahinkar, "Analysis of Performance of Bowlers using Combined Bowling Rate," International Journal of Sports Science and Engineering, vol. 6, no. 3, pp. 1750-9823,2012.

[12].S. Mukherjee,"Quantifying individual performance in Cricket - A network analysis of batsmen and bowlers," Physica A: Statistical Mechanics and its Applications, vol. 393, pp. 624-637, 2014.

[13]. P. Shah,"New performance measure in Cricket," ISOR Journal of Sports and Physical Education, vol. 4, no. 3, pp. 28-30,2017.

[14].Stretch, R. A. (2003). Cricket injuries: a longitudinal study of the nature of injuries to South African cricketers. British Journal of Sports Medicine, 37, pp.250-253

[15].Kimber, A. C., Hansford, A. (1993). A Statistical Analysis of Batting in Cricket. Journal of the Royal Statistical Society,156(3), pp.443-455.

[16].Balasaheb, T., Maman, P., Sandhu J. S. (2008). The impact of visual skills training on batting

performance in cricketers. Serbian Journal of Sports Sciences, 2(1), pp.17-23.