Extraction of Strong and Weak Regions of Cricket Batsman through Text-Commentary Analysis

Muhammad Arslan Rauf
Department of Computer Science
National Textile University
Faisalabad, Pakistan
marslanrauf@hotmail.com

Haseeb Ahmad*
Department of Computer Science
National Textile University
Faisalabad, Pakistan
haseeb_ad@hotmail.com

CM Nadeem Faisal
Department of Computer Science
National Textile University
Faisalabad, Pakistan
nadeem.faisal@ntu.edu.pk

Shahbaz Ahmad
Department of Computer Science
National Textile University
Faisalabad, Pakistan
shahbazsahi@ntu.edu.pk

Muhammad Asif Habib
Department of Computer Science
National Textile University
Faisalabad, Pakistan
dr.m.asif.habib@gmail.com

2020 IEEE 23rd International Multitopic Conference (INMIC) | 978-1-7281-9893-4/20/$31.00 ©2020 IEEE | DOI: 10.1109/INMIC50486.2020.9318089

***Abstract*— Cricket is a famous game in the world where many metrics are introduced and being used to help the coaches and umpires to solve the critical problems. Though different statistics are used to quantify the player's performance based on strike rate, average or for ranking, prediction, and optimal team selection. There is not any effective method to measure the strong and weak regions of a batsman in cricket. In this paper, a text mining method is presented to extract either the strong shot selection points that are frequent for scoring or weak shot regions that seem tough for a batsman to play. Also, a mechanism is put forward to calculate the region-wise strike rate of an individual batsman. To achieve the objectives, the T20 cricket text commentary is being used for this purpose which is available on the espncricinfo website. The proposed method can be helpful for coaches and players to know the strong or weak regions where the batsman feels ease or difficulty to play, respectively. Moreover, the opponent bowlers can also use this method to plan the area where to bowl to each batsman.**

***Keywords*— Sports analytics, Text mining, Region-wise strike rate, T20 cricket**

I. INTRODUCTION

Cricket is a famous sports game in which two teams play and try their best to win the game. In the domain of sports, cricket has gained more importance than other playing games as cricket has now become the second most popular game in the world. Millions of fans throughout the world watch the cricket game. Each playing team is comprised of eleven players including combination of bowlers, batsmen, wicket-keeper and all-rounders. Before the start of each game, the captains of both teams arrive at the ground for the toss. Winner of the toss decides for either batting or bowling first. Three formats of cricket include test cricket, one day International and Twenty20. The twenty20 format is the most recent and famous form of cricket introduced in 2003 [1].

Performance of each player is important, as the result of the match depends on the team runs made by batsmen or the wickets taken by bowlers. The advent of commentary in the form of text has provided an inventiveness to establish models for giving better solutions in the domain of sports. In the game of cricket, textual analysis through commentary could be useful for the player's performance measurement. Player's strength and weakness measured for classification into good, bad and average categories by applying data mining techniques on video frames [2]. In T20, the batsmen tries to hit the ball towards all regions of the ground to make maximum runs. However, each batsman may have some strong and weak regions in terms of making runs and to get out, respectively. Yet, there is no defined procedure to extract the strong and weak regions of each cricket batsman. In this paper, we present a method to extract the strong and weak regions of each batsman concerning pitch delivery and ground regions. Moreover, the region-wise strike rate of each batsman is extracted as well.

A. Ground Regions

Cricket ground has various regions where 11 fielders of the bowling side defend the runs made by opponents. Mostly, the captain decides the fielding placement at best regions where the specific opponent batsman has the chance to play for making maximum runs. There is a circle inside the playing ground based on which fielders are restricted during power-play and other sessions. Different regions of a cricket ground are depicted in Fig.1.

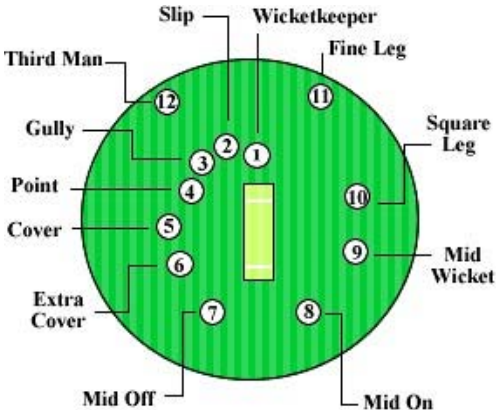

Fig. 1. Different cricket ground regions [3]

978-1-7281-9893-4/20/$31.00 ©2020 IEEE

B. Region-wise Strike Rate

Batsman strike rate (SR), especially in the T20 format is the main attribute to access the performance of a batsman. The overall strike rate of an individual batsman is calculated as the total runs scored divided by the number of balls faced by the batsmen. However, there is no mechanism to calculate the region-wise strike rate rather than the overall strike rate. Region-wise strike rate is measured as the ratio of the count total score made in a specific region (like mid-wicket, mid-on, long-on and long-off) to the number of balls played in that region.

C. Pitch Regions

Cricket pitch is the main area situated at the center of the ground. According to the ICC rules, the size of the pitch is 22 yards in length, where two batsmen run between the wickets for making runs. A pitch includes batting and bowling creases. At the end of the bowling side, umpire stands behind the stump to judge the delivery as legal or illegal. Fig.2 shows the pitch regions where a bowler can deliver the ball. Pitch areas include good length, back of a length, full pitch delivery, short balls and Yorker length balls. Strong and weak pitch regions of cricket batsman are also detected in this work by using these areas of the pitch.

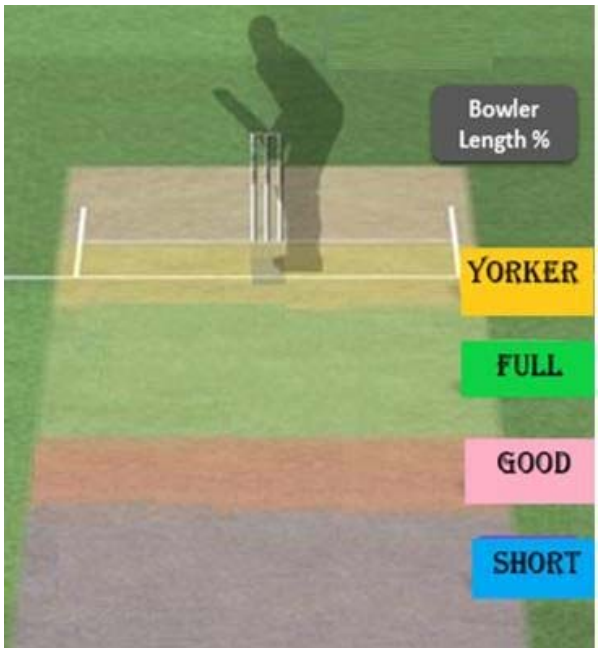

Fig. 2. Pitch having different deliveries areas [4]

The gaps of the strong and weak regions extraction method and the computation of region wise batting strike rates motivated us to present these methods to uplift the analysis in the cricket domain.

The paper is divided into four major sections. In Section I a brief introduction of work is presented. In Section II, related work is provided. The research methodology, data collection and analysis tools are discussed in Section III. The experimental results are depicted in Section IV and Section V consists of the conclusion and future work.

II. RELATED WORK

In the past, different researchers worked in the domain of cricket by using textual, video and numerical data. Text is one of the origins to demonstrate important information. Zhang et al. [5] constructed the news of sports by using text commentary of live events. In this work, the authors summarized the documents to extract the relevant sentence by using a machine learning approach. Live commentary of football is used to extract news events, and "probabilistic sentence selection" algorithm is used to reduce the redundancy. Arif et al. [6] detected the strengths and weaknesses of cricket bowler's by using text commentary. Apriori algorithm is applied on player's data by considering attributes like the home venue or from away playing condition. Jayanth et al. [7] predicted the game outcome by the player's strength and weakness against the specific team. The authors incorporated Support vector machine and players are grouped by K means clustering method.

Nekkanti et al. [8] proposed a mechanism to evaluate the performance of the cricket player's by using Elo rating method to examine the dual between a bowler and batsman. They quantify the performance by measuring the factors like faced and left balls, remaining wickets and skill level of opponent rather than average, strike rate or economy rate. Hussain et al. [9] proposed ranking of teams on the basis of region-wise strength and weakness by implement region-wise team rank and weighted team rank method. Behera et al. [10] explored the strength and weakness of players on the basis of rules defining from textual commentary. They used dimensionality reduction technique to plot a relation between bowler and batsman through biplots.

Fielding plays an important role in winning or losing a match by dropping a catch or good fielding, Perera et al. [11] determine the good player's fielding and run saved by applying random forest method on text commentary data. Sayyed et al. [12] extract the text regions from video by using canny edge detection algorithm. Extraction of cricket stroke played by batsmen from video frames by using camera and applying random forest method for CUTs detection [13]. Sharma et al. [14] proposed a segmentation approach to make a segment of videos by using the text commentary related to that part of the video. The authors applied dynamic programming on video by mapping the segments of a frame with text commentary. Classification methodology is used for measuring the performance of players based on probability. Also, the authors used the clustering method to group the players who have the same playing pattern, to categorize them into a good, bad and average clusters. For this purpose, video frames manually extracted and text annotation stores into the database.

Punjabi et al. [15] presented a prediction model by using naïve bayes to predict the player's performance. It helps to automatic selection of best eleven players against the opposite team in IPL tournament. Perera et al. [16] predicted the optimal line-up of Srilanka cricket team for one day international matches by applying clustering and

association rules mining techniques. To predict the winner of ICC cricket World Cup-2019 which was held in Wales and England, Abdurazzag et al. [17] applied k nearest neighbors algorithm and data reduction technique with Bigdata approach. The authors predicted that India or England can be the winner of the CWC-2019. Passi and Pandey [18] utilized machine learning algorithms to increase the accuracy of prediction in the game of cricket. The authors applied naïve bayes, multiclass support vector machine and decision tree.

To predict the rising stars in the game of cricket, Ahmad et al. [19] used "generative" and "discriminative" machine learning algorithm to distribute both batsmen and bowler according to their scores. The average values of these rising stars are compared with the international cricket council ranking. In one day international matches, preference of one team on the other side is explored recently. Production of bowling and batting at the team level highlight in addition to previous features being used. Daud et al. [20] presented a link fusion method to region-wise ranked the players. They consider inter-type and intra-type base regions relationship for ranking the players. Taj et al. [21] measures the performance of cricket players by using prediction model according to the playing pitch situation. For the selection of team RF algorithm provides the precise result. Ahmad et al. [22] quantify the precedence of cricket team, bowling and batting. Productivity precedence algorithm is presented to extract the precedence of team features. Patil et al. [23] predict the selection of cricket team by applying random forest and decision tree algorithm to build the model.

Albeit of such results, none of the presented works have extracted strong and weak regions of batsmen, moreover, region wise strike rate is also not computed by any of the presented methods.

III. PROPOSED METHODOLOGY

In this paper, we are incorporating a regular expression matching (REM) technique for analyzing the text-commentary of T20 cricket matches to detect the strong and weak regions of a cricket batsman. Moreover, the region-wise strike rate is also calculated through the presented method. The presented research methodology is shown in Fig.3.

Proposed methodology is comprised of the five steps including data collection, data pre-processing, player-wise segmentations and analyzing the region of ground and pitch. Initially, short text-commentary data of each batsman passed through the regular expression for matching the given pattern which contain the attributes such as score, pitch and ground regions. The regular expression is expressed in Equation (1),

$$[0\text{-}9]+(\text{out}|\text{edge})\hat{}[P_r]+[G_r]+ \quad (1)$$

where, P_r = {short-length|good-length|full-length|yorker} and G_r = {midwicket|third-man|square-leg|long-on|long-off|fine-leg}. Consider the following text-commentary line,

> *3.2, Ali to **Babar Azam**, **FOUR**, **short of a length**, but a little wide, enough for Babar to stand tall and punch it at **point***

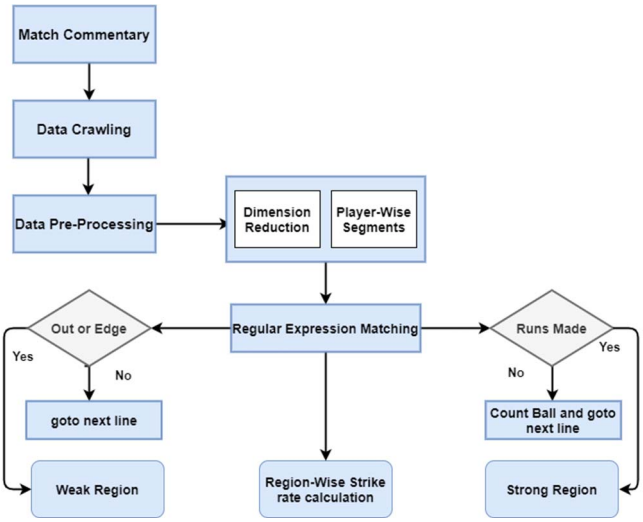

Fig. 3. Proposed Methodology

Batsman name (Barbar Azam), outcome of the ball (FOUR), pitch region (Short of a length) and short area (point) are described in the given commentary. These features are used for matching with regular expression to calculate the score or measure the out or edge at each region where the pattern match.

A. Data Collection and Analysis Tools

We initially crawled data from the sports website ESPN cricinfo by using scraping tools, python with spyder interface. Also, different library packages including beautiful soup, URL request and selenium for web automation are used which helped to extract text commentary dynamically by preventing the effect of page scrolling. This textual annotation of cricket strokes contains every action that occurred during the real-time event.

Inning by inning list of data includes 2, 13,000 text commentary lines that were imported in the CSV file. Dataset has included all the international players which played T20 cricket during the session 2008 to 2019.

B. Data Pre-processing

After collecting a large amount of data, the preprocessing is done to reduce the dimensionality of the data. Since it is not easy to extract relevant information from the large data set. Efficiency can be improved by reducing the amount of irrelevant data. Trimming the data by removing the extraneous lines of commentary such as "end of over information" at the end of each overwrite down on cricket website. Also, the unimportant discussion between the commentators is eliminated. We preprocessed text commentary data by removing the stop words which contain punctuation marks, and prepositions by using python language with NLTK libraries including stop words, Stemming and WordNet lemmatizer to make the text consistency and well-organized form. Aftewards, data separated by producing the player-wise csv files. Each file contains text-commentary of individual batsman.

C. Strength and Weakness Analysis

Proposed regular expression algorithm for measuring the strong and weak regions of cricket batsmen is presented in the given REM approach.

Proposed REM Approach

INPUT: the set of source files S_{files}
OUTPUT: batsmen strong S_r, weak W_k region
F_i : i^{th} line in file
G_r : ground region
P_r : pitch region
region_score ← **Ø** // score 1,2,3,4,6
balls ← **Ø** // balls played in specific region
for each line F_i in S_{files} **do**
  if (re.search = = P_r) && (re.search = = score)
    if (re.search = = G_r)
    balls = lines F_i
    region_score += balls.score
    inc i
    total_run = sum (region_score)
  for each G_r in file
    max = total_run
    if (G_r.total_run >= max)
      $S_r = G_r$
  else
    if (re.search = = P_r) && (re.search = = edge || out)
    for each G_r in file
    count = lines F_i
    inc i
    max = len (count)
    if (G_r.len(count) >= max)
      $W_k = G_r$

In the proposed mechanism, each batsman source file is read line by line and each commentary line is parsed into tokens based on word_tokenize. The token is passed through the regular expression i.e. score, pitch and ground region for pattern matching. For each batsman, pitch area where a bowler can bowl like a good-length ball, full length-ball or shorter length ball are mapped with given commentary lines. After matching with these pitch regions, the model checks that how many scores are made in that area of the pitch and in which ground region the batsman played. Otherwise, the token is simply ignored and the next line of commentary is fetched. After calculating scores at each region, a comparison is made with all regions of ground along with pitch area. If batsmen make the highest score by playing at specific ground region among all ground regions then it is considered as the strong playing area of cricket batsmen. Similarly, the weak region is that area of ground and pitch where a batsman gets out for the maximum time or feels it difficult to play.

D. Region-Wise Strike rate Calculation

Since the strike rate is the most effective feature the strike rate in the short format T20 of the cricket game. Two batsmen having the same number of runs, may have different strike rates. Batsman with higher strike rate is considered to be better. Moreover, a batsman is considered better if he has better strike rates in most regions of the ground. The region-wise strike calculation is a new task as, before this research, an overall strike rate of batsmen were already examined. Overall strike rate were not helpful to judge the performance of batsmen by region-wise. We present a method which calculates the region-wise strike rate of cricket Batsman. The region-wise strike rate is measured by the statistical formula as shown in Equation (2) as follows:

$$SR_{br} = R_{br}/BP_{br} \quad (2)$$

Strike rate is calculated as the total number of the runs made in specific regions like mid-wicket, long-on or long-off divided by the total number of balls played in the underlying region. Here, SR_{br} represents the strike rate of batsman b in region r. R_{br} depicts the runs made by batsman b in r. BP_{br} presents the balls played by batsman b in region r.

IV. EXPERIMENTAL RESULTS

In this work, a text-commentary dataset of seventy T20 cricket batsmen is used to perform experiments using regular expression based matching system. The dataset contains all the T20 international matches played from the session 2008 to 2019. Cricket batsmen including from all the international playing who played twenty over game.

A. Region-based Batsmen Strength and Weakness

To measure the strong and weak regions of cricket batsmen, top ten T20 cricket batsmen of ICC ranked in September 2019 are examined. Table 1 represents the strong ground regions of batsmen when they played according to the pitch length balls. Babar Azam makes a higher score on Midwicket and Cover point when any bowler delivers the ball as the back of a length and full-length balls. This frequent pattern represents the strength of Babar Azam. Maxwell have strong Midwicket region when he played mostly from the back of a length and slower pitch areas. Similarly, other batsmen strong regions are depicted in Table 1.

Table 1: Strong regions of T20 Cricket Batsmen w.r.t. Pitch

Batsman Name	Team	Region (Pitch)	Region (Ground)
Babar Azam	Pakistan	Back of length and Fuller	Cover and Midwicket
Maxwell	Australia	Back of length and Slower	Midwicket
Munro	New Zealand	Full-length ball	Long-on
Finch	Australia	Back of length	Square Leg
Ross Taylor	New Zealand	Full-and Short balls	Midwicket and Square- leg
Rohit Sharma	India	Shorter and Slower balls	Square leg and Midwicket
Kane Williamson	New Zealand	Full-length balls	Cover Point
Fakhar Zaman	Pakistan	Back of length balls	Cover point and Square leg
V Kohli	India	Full and Back of length	Midwicket
De Kock	South Africa	Back of length	Point and Midwicket

Table 2 represents the weak ground regions of batsmen when they played according to the pitch length balls. Babar Azam maximum time got out or made edges at point when any bowler delivered the ball at the shorter and full-length pitch regions. This frequent pattern represents the weakness of Babar Azam at a specific area. Similarly, other batsmen weak regions are depicted in Table 2.

Table 2: Weak regions of T20 Cricket Batsmen w.r.t. Pitch

Batsman Name	Team	Region (Pitch)	Region (Ground)
Babar Azam	Pakistan	Fuller and Shorter balls	Point
Maxwell	Australia	Good-length balls	Third man
Munro	New Zealand	Full length ball	Square leg and Third man
Finch	Australia	Full and good length balls	Third-man and Point
Ross Taylor	New Zealand	Full and short balls	Midwicket and Point
Rohit Sharma	India	Full and good length balls	Long-on
Kane Williamson	New Zealand	Shorter and full length	Midwicket
Fakhar Zaman	Pakistan	Back of length balls	Cover and Midwicket
V Kohli	India	Full and good length balls	Point and Slips
De Kock	South Africa	Good length and Short ball	Point and Third-man

The underlying work presents a method to extract the strong and weak regions of cricket batsmen by considering the pitch areas with ground regions. It shows that, for particular pitch deliveries, batsmen make more score in a specific area of field position or get out or feels difficult to play. Result describes that Babar Azam has Midwicket and Cover point as strong regions, while third man and Point are extracted as weak regions of the ground by playing on every type of pitch delivery. Moreover, batsmen strong and weak ground or fielding regions are comprised regardless of pitch deliveries. A batsmen played frequently on the specific region by playing on all pitch deliveries like Short ball, Full-length balls, Back of length or Yorker. By considering all the pitch length-balls, overall strong and weak ground regions of batsmen are depicted in Table 3.

B. Region-wise Strike Rate of T20 Batsmen

To calculate the region-wise strike rate, twelve regions of cricket ground are included which are Midwicket, Mid-off, Mid-on, Square-leg, Fine-leg, Cover point, Extra cover, Point, Long-on, Long-off, Straight and Third-man. Table 4 represents the cricket batsmen strike rates at each region. In short format of game, strike rate also plays an important role to judge the performance of batsmen.

Table 3: Region-based Strong and Weak Regions

Batsman Name	Team	Strong Regions	Weak Region
Babar Azam	Pakistan	Midwicket and Cover point	Third-man and point
Maxwell	Australia	Midwicket	Point and deep Midwicket
Munro	New Zealand	Deep midwicket	Long-on
Finch	Australia	Cover point and Midwicket	Square leg
Ross Taylor	New Zealand	Long-on and Midwicket	Midwicket and square leg
Rohit Sharma	India	Midwicket	Third man
Kane Williamson	New Zealand	Square leg and Midwicket	Midwicket and cover
Fakhar Zaman	Pakistan	Midwicket and Cover point	Third man
V Kohli	India	Long-on and midwicket	Point
De Kock	South Africa	Midwicket	Cover and Midwicket

Table 4: Region-wise Strike Rate of T20 Cricket Batsmen

Batsman Name	Team	Midwicket	Mid-on	Mid-off	Square-leg	Fine leg	Cover point	Extra cover	Point	Long on	Long off	Straight	Third man
Babar Azam	Pakistan	150	70	121	122	181	123	100	152	151	158	112	157
Maxwell	Australia	198	150	214	213	291	175	150	231	223	161	150	195
Munro	New Zealand	227	54	172	234	153	193	200	154	150	242	145	250
Finch	Australia	220	165	188	200	241	154	227	150	247	230	148	170
Ross Taylor	New Zealand	162	100	58	195	203	116	151	135	222	164	79	173
Rohit Sharma	India	166	92	164	163	274	104	109	163	196	251	118	156
Williamson	New Zealand	154	105	163	173	182	129	231	164	123	188	144	130
Fakhar	Pakistan	177	110	142	206	158	139	181	179	260	184	155	200
V Kohli	India	161	140	145	163	225	168	244	152	180	198	127	179
De Kock	South Africa	201	150	134	136	141	150	300	129	192	164	87	138

V. CONCLUSION AND FUTURE WORK

In this paper, we extracted the strengths and weaknesses of cricket batsmen in the context of text-commentary. Moreover, the proposed method also determines the strike rate of cricket batsmen by regions like midwicket, mid-on, mid-off, long-on, long-off, deep midwicket, and square leg, etc. In the proposed method, data set is collected from espncricinfo sports website. The data set includes the text commentary of International T20 playing matches from 2008 to 2019. From these lines of commentary, player's strong and weak regions measured by considering pitch and ground areas. The player's strength and weakness analyses are necessary to judge their performance. It will be helpful to the coaching staff to improve the weakness of players. Also, to make a better strike rate where they feel difficulty to play and make the score. The opposition team can also use this mechanism to know the weakness of batsmen and plan according to the opponent weakness. Region-wise strike rate can be used for setting field position during the match.

This approach can be applied for the other format of the game including test and one day international matches. For test matches, other parameters can also be included to show their ability to play longer innings instead of strike rate. Also, this work can be applied to calculate player-wise (against specific bowler) strike rate.

REFERENCES

[1] Manage, Ananda BW, and Stephen M. Scariano. "An introductory application of principal components to cricket data." Journal of Statistics Education 21, no. 3 (2013).

[2] Sharma, Rahul Anand, K. Pramod Sankar, and C. V. Jawahar. "Fine-grain annotation of cricket videos." In 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 421-425. IEEE, 2015.

[3] Ground regions, "sportskeeda," [Online]. Available: https://statics.sportskeeda.com/wpcontent/uploads/ 2013/12/adelaide-pitch-map-2030773.png, [Accessed 29 December 2019].

[4] Cricket pitch, "wikipedia," [Online]. Available: https://en.wikipedia.org/wiki/Cricket_pitch#/media/File:Cricket_ pitch.svg. [Accessed 29 12 2019].

[5] Zhang, Jianmin, Jin-ge Yao, and Xiaojun Wan. "Towards constructing sports news from live text commentary." In Proceedings of the 54th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers), pp. 13611371. 2016.

[6] Arif, Subhan, Muhammad Umair, Syed Muhammad Kumail Naqvi, Ahmad Ikram, and Anam Ikram. "Detection of bowler's strong and weak area in cricket through commentary." In Proceedings of the 2nd International Conference on Future Networks and Distributed Systems, pp. 1-14. 2018.

[7] Jayanth, Sandesh Bananki, Akas Anthony, Gududuru Abhilasha, Noorni Shaik, and Gowri Srinivasa. "A team recommendation system and outcome prediction for the game of cricket." Journal of Sports Analytics 4, no. 4 (2018): 263-273.

[8] Nekkanti, Yamini, and Dibyojyoti Bhattacharjee. "Novel Performance Metrics to Evaluate the Duel between a Batsman and a Bowler." Management and Labour Studies 45, no. 2 (2020): 201-211.

[9] Hussain, Akbar, Yan Qiang, M. Abdul Qadoos Bilal, Kun Wu, Zijuan Zhao, and Bilal Ahmed. "Region-wise ranking for one-day international (ODI) cricket teams." machine learning 10, no. 10 (2019).

[10] Behera, Swarup Ranjan, Parag Agrawal, Amit Awekar, and Vijaya Saradhi Vedula. "Mining Strengths and Weaknesses of Cricket Players Using Short Text Commentary." In 2019 18th IEEE International Conference On Machine Learning And Applications (ICMLA), pp. 673-679. IEEE, 2019.

[11] Perera, Harsha, Jack Davis, and Tim B. Swartz. "Assessing the impact of fielding in Twenty20 cricket." Journal of the Operational Research Society 69, no. 8 (2018): 1335-1343.

[12] K. Tabish Sayyed, "Super Imposed Method for Text Extraction in a Sports Video," International Journal on Future Revolution in Computer Science & Communication Engineering, vol. 4, no. 3, pp. 332-337, 2018.

[13] Gupta, Arpan. "Cricket stroke extraction: Towards creation of a large-scale cricket actions data-set." arXiv preprint arXiv: 1901.03107 (2019).

[14] Sharma, Rahul Anand, K. Pramod Sankar, and C. V. Jawahar. "Fine-grain annotation of cricket videos." In 2015 3rd IAPR Asian Conference on Pattern Recognition (ACPR), pp. 421-425. IEEE, 2015.

[15] Punjabi, Vipul, Rohit Chaudhari, Devendra Pal, Kunal Nhavi, Nikhil Shimpi, and Harshal Joshi. "A Survey on Team Selection in Game of Cricket using Machine Learning." (2019).

[16] Perera, Harsha, Jack Davis, and Tim B. Swartz. "Optimal lineups in Twenty20 cricket." Journal of Statistical Computation and Simulation 86, no. 14 (2016): 2888-2900.

[17] Abdurazzag A Aburas, Muhammed Mehtab and Yusuf Mehtab, "Cricket World Cup Predictions Using KNN Intelligent Bigdata Approach," in ACM, Charleston, SC, USA, 2018.

[18] N. Kalpdrum Passi, "Increased prediction accuracy in the game of cricket using machine learning," International Journal of Data Mining & Knowledge Management Process (IJDKP), vol. 2, no. 2, pp. 1-18, March 2018.

[19] Ahmad, Haseeb, Ali Daud, Licheng Wang, Haibo Hong, Hussain Dawood, and Yixian Yang. "Prediction of rising stars in the game of cricket." IEEE Access 5 (2017): 4104-4124.

[20] Daud, Ali, Akbar Hussain, Rabeeh Ayaz Abbasi, Naif Radi Aljohani, Tehmina Amjad, and Hassan Dawood. "Region-wise Ranking of Sports Players based on Link Fusion." In Companion Proceedings of the The Web Conference 2018, pp. 259-266.

[21] Taj, S. Ghouhar, T. Bhavana, E. V. Pravallika, P. Saileswar Reddy, and G. Swarnalatha. "Machine Learning Model for Predicting the Performance Analysis of Cricket Players'." Sustainable Humanosphere 16, no. 2 (2020): 17-23.

[22] Ahmad, Haseeb, Ali Daud, Licheng Wang, Ibrar Ahmad, Muhammad Hafeez, and Yixian Yang. "Quantifying team precedence in the game of cricket." Cluster Computing 21, no. 1 (2018): 523-537.

[23] Patil, Nilesh, Bevan H. Sequeira, Neil N. Gonsalves, and Abhishek A. Singh. "Cricket Team Prediction Using Machine Learning Techniques." Available at SSRN 3572740 (2020)