Journal of Applied Science, Engineering, Technology and Management
Vol. 1, Issue 1, June 2023 pp. 21 - 27

Prediction of Right Bowlers for Death Overs in Cricket

Lakshmi Shaju
Department of Computer Science and Engineering, Federal Institute of Science and Technology, Angamaly, Kerala, India
lakshmishaju300@gmail.com

Rahul K Sajith
Department of Computer Science and Engineering, Federal Institute of Science and Technology, Angamaly, Kerala, India
rahulksajith2016@gmail.com

Reshma N P
Department of Computer Science and Engineering, Federal Institute of Science and Technology, Angamaly, Kerala, India
reshmapoulose998@gmail.com

Vinayak N
Department of Computer Science and Engineering, Federal Institute of Science and Technology, Angamaly, Kerala, India
vinayaknmenon@gmail.com

Paul P Mathai
Department of Computer Science and Engineering, Federal Institute of Science and Technology, Angamaly, Kerala, India
paulpmathai@fisat.ac.in

***Abstract*—Predicting the right bowlers for the death overs in cricket is crucial for the success of a team, as these overs can often determine the outcome of a match. Death overs refer to the final overs of an innings, which are usually considered to be the most crucial because they can determine the result of the match. The death overs are typically the last 5-6 overs of an innings, and are known for being high-pressure situations due to the need to score runs or take wickets. Batting team has always an advantage in winning the match, by selecting the right bowlers to bowl in death overs, it would create an edge for the bowling team to win the match. Selecting the right bowlers for death overs requires an in-depth understanding of the strengths and weaknesses of different bowlers and the ability to analyze various factors such as the pitch conditions, the opposition team, and the current game situation. Some of the key factors that can be used to predict the right bowlers for death overs include the bowler's past performance in similar situations, their ability to bowl under pressure, and their ability to take wickets, their ability to bowl Yorkers and slower balls. In order to accurately predict the right bowlers for death overs, it is important to consider all of these factors and use statistical analysis and machine learning techniques to build models that can accurately predict their performance. This analysis can then be used to build models that can accurately predict which bowlers are likely to perform well in the death overs based on these factors.**

Keywords*— *Classifier, Decision Tree, Pre-processing, Random Forest, Regular expression, SVM.

I. INTRODUCTION

Cricket is a popular sport played between two teams, each consisting of eleven players. It is primarily played in countries such as England, Australia, India, Pakistan, South Africa, and the West Indies, but its popularity has spread to various parts of the world. Cricket is known for its rich history, traditions, and passionate fan following. The objective of cricket is to score more runs than the opposing team while taking the wickets of their batsmen. The game is played on a large oval-shaped field, known as the cricket ground, which features a 22-yard-long pitch at the centre. The pitch has two sets of three wooden stumps, called wickets, at each end, with two small bails resting on top of the stumps. A cricket match consists of two innings, with each team having a chance to bat and bowl. The team batting attempts to score runs by striking the ball bowled by the opposing team's bowler and then running between the wickets or hitting boundaries, while the fielding team tries to dismiss the batsmen and limit the number of runs scored. The two main forms of cricket are Test cricket and limited-overs cricket. Test cricket is the longest format, played over five days with each team having two innings. It is considered the pinnacle of the sport and requires endurance, skill, and strategic planning. Limited-overs cricket includes One Day Internationals (ODIs) and Twenty20 Internationals (T20Is). In ODIs, each team has 50 overs to bat and bowl, while T20Is are even shorter, with each team facing 20 overs. Overall, cricket is a captivating sport that combines skill, strategy, and teamwork. Its unique blend of tradition and innovation continues to captivate audiences around the world, making it one of the most beloved sports on the planet.

It was intended to create a fast-paced game that would appeal to both on-the-ground fans and broadcast watchers. The pastime has spread throughout the cricket world. Most international tours include at least one Twenty20 encounter, and every Test-playing country has a local cup tournament. Following that, the Big Bash League, Bangladesh Premier League, Pakistan Super League, Caribbean Premier League, and Afghanistan Premier League all used identical formulas and were well-liked by supporters. The Women's Big Bash League was established by Cricket Australia in 2015, and the Kia Super League began in England and Wales in 2016. The Mzansi Super League of South Africa began in 2018. Numerous T20 competitions follow the general pattern of a group stage followed by a Page playoff system among the top four teams in which:

- The first- and second-placed teams in the group stage encounter, with the winner progressing to the final.

Received for review: 07-04-2023
Accepted for publication: 06-06-2023
Published: 16-06-2023

- The third- and fourth-place teams square off, with the loser ousted; and
- the two teams that have not yet progressed to the final after the previous two matches are played fight for the second spot in the final.

In the Big Bash League, an extra encounter is contested to decide whether the fourth or fifth-placed club will qualify for the top four. We acquire the necessary data from websites such as Cricinfo, Cricbuzz, and Cricmetric, which will provide us with our information. The data is then assembled in order to extract the required traits as effectively as possible. Bowling average, wickets gained, economy, and dot ball average are used to evaluate the bowler against a specific batter, whereas hitting strike rate and batting average are used to evaluate the striker. These acquired data are preprocessed so that the ml model can predict correctly. We suggest using Decision Tree with tagging, Random Forest, and SVM as ML models. As a consequence, for each batter, a similar bowler from the opposing side will be expected. This improves the hitting team's odds of winning.

II. Related works

The work [1] aims to develop a system that can predict cricketers' performance in different tournaments and assist with squad selection. A lot of historical data on cricket players' competition performance is gathered, along with details on other important factors including player age, fitness level, playing conditions, and team composition. Technology Features The relevant elements, such as player performance measurements, team performance metrics, match conditions, player attributes, and historical trends, are then extracted from the obtained data. The batting average, bowling average, team balance, and (if relevant) budgetary constraints should all be considered when choosing a squad. With the help of the trained model, it is possible to forecast cricket players' performance in upcoming competitions, including their batting and bowling averages and other important variables. Based on expected performance and specified objectives, multi-objective optimization techniques are used to choose the ideal squad. To determine the success of the system, the performance of chosen squads is compared to test results or prior performance.

In order to identify a batsman's strong and weak areas, this method [2] focuses on textual commentary in cricket matches. Information on cricket games' text commentary is gathered. This provides ball-by-ball summaries of the game, along with information on the batter's choice of shot, field placements, and results. The commentary is examined to determine the batsman's style of strokes. The tendencies and preferences of the batsman can be discovered by examining the shot selection. The results of each shot the batter takes are examined. This provides data on the number of runs, boundaries, wickets, or dot balls produced by the stroke. The efficacy of the batsman's strokes in various parts of the field may be ascertained by examining the results. Each shot the batsman plays is assigned to the appropriate section of the field based on the shot selection and results. The strong and weak areas of the batsman may be identified by combining the data on shot choices and results mapped to particular field sections. The analysis's findings are shown graphically to help comprehend the batsman's strong and weak areas. Heat maps, charts, or other graphical representations that emphasize the batsman's success in various fielding locations can be used to do this. The batsman's identified strong and weak areas can be assessed using historical data and compared to those of other batters. The batsman's playing style, strengths, shortcomings, and opportunities for development can be better understood through this examination. In order to identify and portray the batsman's strong and weak areas, the project calls for the use of natural language processing (NLP) techniques to extract pertinent information from the textual feedback.

In order to pick a cricket team that concurrently optimizes several goals, optimization is a method that makes use of evolutionary algorithms. These goals may involve improving team balance, batting efficiency, bowling efficiency, fielding prowess, and other pertinent factors. Each prospective team is modelled as a possible solution, generally as a binary string or a numerical vector. To start the evolutionary process, a population of randomly selected or predetermined candidate solutions (teams) is formed. Based on the specified objectives, each applicant team is assessed within the population [3]. This entails evaluating each player's performance as well as that of the entire team using a variety of indicators, including batting average, bowling economy, fielding statistics, and team balance. The most potential applicant teams are chosen based on their fitness ratings using a selection process, such as a tournament or roulette wheel. This process makes sure that teams who do better are more likely to be chosen for the following generation. Crossover and mutation are two examples of genetic operators that are used to the chosen teams to produce progeny. While mutation brings arbitrary alterations to the genetic make-up of individual teams, crossover combines the genetic composition of two parent teams to generate new teams. The most physically fit candidate teams are chosen to make up the following generation, both from the parent population and the offspring population. The better-performing teams are more likely to survive and advance to succeeding generations thanks to this selection process. The ideal team(s) that accomplish the intended goals are thought to be the top-performing candidate team(s) from the last generation. The cricket team's ultimate selection is made up of these teams. This strategy finds a collection of trade-off solutions that balance several criteria by using evolutionary multi-objective optimization, which enables simultaneous evaluation of many objectives.

III. Problem statement

To select the best bowler from the given set of bowlers present in the team to bowl against the batting pair or set the batter during the death overs of the game using stats such as economy, no of times wicket taken, bowling average, batting average against the bowler, dot ball percentage, etc so as to reduce the bias towards the batting team. So, we aim to reduce the bias by implementing two different models, one using the classifier algorithms and the next using NLP. In order to identify the best bowler to bowl against a specific batter, various factors need to be taken into consideration. One important factor is the economy rate of the bowlers against the batter. This metric measures the number of runs a bowler gives up per over, when bowling to the specific batter. A lower economy rate indicates that the bowler is more effective at limiting the runs scored by the batter. Another important

factor is the number of times the bowler has taken the batter's wicket. A bowler who has a higher number of wickets against a specific batter is more likely to be successful at getting the batter out. Additionally, the dot ball percentage of a bowler against a specific batter can also be taken into consideration. This metric measures the number of balls bowled by the bowler in which the batter does not score a run. A higher dot ball percentage indicates that the bowler is more effective at restricting the batter from scoring runs. By analyzing these factors and comparing the performance of different bowlers against a specific batter, it is possible to identify the best bowler to bowl against that batter.

IV. SYSTEM ARCHITECTURE

Fig. 1 Shows the system architecture. It includes data collection and preprocessing, training, analysis and prediction.

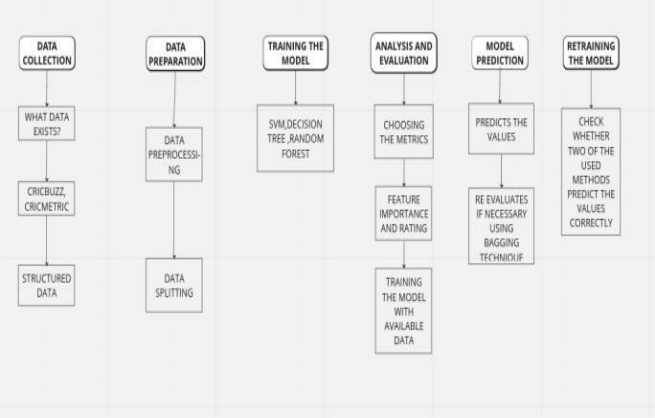

Fig. 1. Block Diagram (Using Classifier Algorithm

V. DESIGN METHODOLOGIES

Selecting the best bowler for the death overs in T20I cricket is of utmost importance, and it is equally crucial to reduce bias towards the batting team. During the death overs, the batting team aims to score as many runs as possible. A skilled death bowler can restrict the opposition's scoring rate by executing Yorkers, slower deliveries, and variations effectively. This ability to control the run flow can significantly impact the final outcome of the match. The Objective to create models one using classifier algorithms and the other one with NLP which will predict the bowler against a particular batsman.

B. Methodology

Using Classifier Algorithms: The data collection for the model using classifier algorithms were done manually. The data was taken from site namely cricmetric where the relevant data for players from 10 countries was collected. and analyzed. The data included player statistics such as batting average, bowling average, strike rate, economy rate, and fielding performance. The data was then pre-processed to remove any missing values and outliers. Feature selection techniques were applied to select the most relevant features for the model. The selected features were then used to train various classifier algorithms such as decision trees, random forests, support vector machines. The performance of each algorithm was evaluated using metrics such as accuracy, precision, recall, and F1 score. The results showed that the random forest algorithm outperformed the other algorithms. Random Forest algorithm m is a machine learning technique that utilizes decision trees to classify data. It is particularly useful for handling large data-sets with multiple variables, as it can identify the most important features for classification. The algorithm works by creating a multitude of decision trees and then combining their predictions to produce a final result. This approach helps to reduce over-fitting and improve the accuracy of the model. In addition, random forests are robust to outliers and missing data, making them a reliable choice for real-world applications. Overall, the model proved to be effective in predicting player performance based on the statistics and can be used by cricket teams for player selection and strategy planning.

Using NLP: The dataset for this NLP model was gathered from Kaggle and included three years of cricket commentary from both international and domestic competitions. The dataset contains information about all game formats. The dataset included data from 2017 to post-covid periods Because the data was in a mixed format, we had to separate the t20 international matches from the others. We did the preprocessing on the computers remotely. We wanted to determine a batsman's vulnerability by identifying his weak ground and pitch locations, and then identifying the bowlers he struggled against using data from t20i matches from the top ten t20i ranking countries. Weak Ground zone: To determine the weak ground zone, we created a dictionary of frequent phrases used by commentators to specify a location, such as 'midwicket', 'mid-wicket', and 'mid-wicket' for the mid-wicket region, and the same for the other regions. To find the weak pitch region, we created a dictionary of common terms used by commentators to specify a pitch region such as 'short and', 'short delivery', 'short of a length', 'ball short', 'short ball', 'short-ball', 'short-length', 'short length', 'shorter', 'shortish' for the short ball pitch region and the same for other regions. Bowler The hitter Is Weak Against: If a bowler consistently affects the count of a hitter's weak pitch and ground location, he becomes a bowler the batter is weak against. We utilize the spacy NLP model to count the appearances of the top three weak regions and top three weak pitch regions, and then we map the top three bowlers against whom the batter struggled based on these two criteria by comparing the dates of appearances in the dataset rows.

VI. IMPLEMENTATION

A. Using Spacy NLP

SpaCy is a highly capable Natural Language Processing (NLP) toolkit that offers an extensive range of tools for text processing, analysis, and machine learning. With its robust functionality and efficient performance, SpaCy has become a preferred option for a wide array of NLP tasks. Its features include reliable tokenization, accurate Part-of-Speech (POS)

tagging, Named Entity Recognition (NER), dependency parsing, lemmatization, and text classification. SpaCy can be seamlessly integrated with popular Python libraries and frameworks such as scikit-learn, TensorFlow, and PyTorch, enabling the development of end-to-end NLP pipelines. It also supports multiple languages and provides pre-trained models.

The SpaCy community actively contributes to its development, offering extensive documentation, tutorials, and examples to facilitate rapid adoption. By leveraging SpaCy's capabilities, you can enhance NLP projects, automate text analysis workflows, and extract valuable insights from vast volumes of textual data.

1) Dataset: The dataset for this NLP model was obtained from Kaggle which contained 3-year cricket commentary of various cricket matches, both international and domestic leagues. The dataset contained data regarding all formats of the game. The dataset contained information from 2017 to post-covid times.

2) Data Preprocessing: The dataset had data regarding test matches, ODI matches and also t20 matches. It also had data of various domestic fantasy leagues. All the data were in a mixed format, so we had to separate the t20 international matches from all. We performed the preprocessing locally on our computers. We used basic python modules like OS and shutil for this. All the datasets that had rows that were almost equivalent of that of a t20 match was copied to another directory.

   Then we separated the csv's to that of different countries again by running a python script and checking the country name in the specific attribute column of the dataset. This was not done manually as it would have been a very tedious task. After that we, made folders for storing the data of t20i matches of the top 10 t20i ranking countries and then further proceeded with the project.

3) Weakness Analysis: We intended to find the weakness of a batter by finding his weak ground regions and weak pitch regions, and then finding the bowlers whom he struggled against.

   Weak Ground Region: To find the weak ground region, we made a dictionary of common terms used by commentators to specify a region like 'midwicket', 'mid-wicket' and 'mid-wicket' for mid-wicket region and the same for other regions.

   Weak Pitch Region: To find the weak pitch region, we made a dictionary of common terms used by commentators to specify a pitch region like 'short and', 'short delivery', 'short of a length', 'ball short', 'short ball', 'short-ball', 'short-length', 'short length', 'shorter', 'shortish' for short ball pitch region and the same for other regions.

   Bowler The Batter Is Weak Against: If a bowler has constantly affected in increasing the count of the weak pitch and ground region of a batter, then he becomes a bowler the batter is weak against. We use spacy NLP model and find out the count of the appearances of the top 3 weak region and top 3 weak pitch region and map the top 3 bowlers against whom the batter struggled using these two criteria by comparing the times of appearances in the dataset rows.

4) Libraries/Applications:

   OS and shutil: OS and shutil is used for locally manipulating the csv datasets to sort out the specific datasets required for performing the weakness analysis.

   SpaCy: SpaCy was used for performing the weakness analysis by reading the commentary dataset and finding the weak pitch region, weak ground region and by using those data, finding the top 3 bowlers the batter is weak against.

A. Classifier Models

Different Algorithms like Decision Tree, Random Forest are used in order to predict the right bowler against a batsman during death overs. A decision tree is a machine learning algorithm that can be used for both classification and regression tasks. It is a flowchart-like model where each internal node represents a feature, each branch represents a decision based on that feature, and each leaf node represents the outcome or prediction. The decision tree algorithm recursively partitions the data based on different features, aiming to create branches that maximize the separation between different classes or minimize the variance within each branch. This process is guided by various metrics, such as information gain or Gini impurity, to determine the best feature and threshold for splitting.

Support Vector Machines (SVM) is a supervised machine learning algorithm commonly used for classification and regression tasks. SVMs have been widely applied in various domains, including finance, healthcare, image recognition, and natural language processing. SVM is primarily known for its classification capabilities. It constructs a hyperplane or a set of hyperplanes in a high-dimensional feature space to separate data points into different classes.

Random Forest has several advantages that make it a popular and effective machine learning algorithm for various prediction tasks. Random Forest is known for its robustness against overfitting. Overfitting occurs when a model performs well on the training data but fails to generalize to new, unseen data. Random Forest mitigates overfitting by creating multiple decision trees and aggregating their predictions, reducing the impact of individual noisy or biased trees.

Dataset: Data of players from 10 countries were acquired to obtain accurate and reliable data. It was analyzed using various statistical methods to identify trends and patterns. There were significant differences in the performance of players from different countries, with some countries consistently producing top-performing players.

There were certain factors, such as age and experience, that were strongly correlated with player performance. Armed with this knowledge, we were able to develop a predictive model that could accurately forecast the performance of players based on their demographic and performance data.

This model proved to be highly effective in predicting the outcomes of matches and tournaments, and helped us to make informed decisions about which players to recruit for our team.

B. Data Preprocessing

The information was prepared for the classifier algorithms' training. This could entail coding categorical variables, handling missing values, and normalizing or scaling numerical features. We had made sure that the data was in a format that would be appropriate for the next steps. process of machine learning. Following the preparation of the data, we chose and trained our classifier algorithms. This required examining various models and deciding, based on performance metrics, which one was the best. Then, in order to maximize the accuracy of our chosen model, we adjusted its hyperparameters. After training, we ran our model against a different set of data to assess how well it performed in practical situations.

VII. LIBRARIES/APPLICATIONS

a) SpaCy

The spaCy package is a popular and powerful natural language processing (NLP) library in Python. It is designed to be efficient, fast, and user-friendly, making it a valuable tool for various NLP tasks, such as tokenization, part-of-speech tagging, named entity recognition, syntactic dependency parsing, and more.

b) Pandas

Pandas is defined as an open-source library that provides high-performance data manipulation in Python. Data analysis requires lots of processing, such as restructuring, cleaning or merging, etc. There are different tools are available for fast data processing, such as Numpy, Scipy, Cython, and Panda. But we prefer Pandas because working with Pandas is fast, simple and more expressive than other tools.

c) Scikit-Learn

Scikit-learn (Sklearn) is the most useful and robust library for machine learning in Python. It provides a selection of efficient tools for machine learning and statistical modeling including classification, regression, clustering and dimensionality reduction via a consistence interface in Python.

d) Streamlit

Streamlit is an open-source app framework in Python language. It helps us create web apps for data science and machine learning in a short time. It is compatible with major Python libraries such as scikit-learn, Keras, PyTorch, SymPy(latex), NumPy, pandas, Matplotlib etc.

e) RE

A regular expression (or RE) specifies a set of strings that matches it; the functions in this module let you check if a particular string matches a given regular expression (or if a given regular expression matches a particular string, which comes down to the same thing).

f) OS

The OS module in Python provides functions for creating and removing a directory (folder), fetching its contents, changing and identifying the current directory, etc. Need to import the OS module to interact with the underlying operating system.

g) CSV

The csv package in Python is a built-in module that provides functionality for reading and writing Comma-Separated Values (CSV) files. CSV files are a common way of storing and exchanging data, where each record is represented as a line and fields within the record are separated by commas.

VIII. RESULTS

In the classifier feature, according to the dataset we have acquired manually bowler for the particular batsman is predicted. Out of all the algorithms used, random forest has the highest accuracy.

Fig. 2. Result of SVM Algorithm

Fig. 3. Result of Decision Tree Algorithm

Fig. 4. Result of Random Forest Algorithm

In the NLP feature, Data is taken from a Kaggle commentary dataset. This will search for the bowler, pitch regions and ground regions corresponding to the difficult words, pitch region words and ground region words declared respectively using spacy NLP.

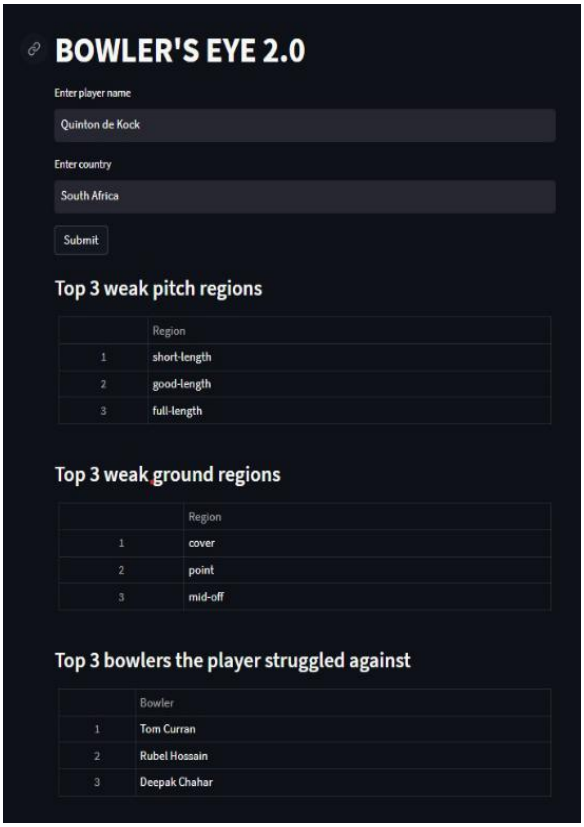

Fig. 5. With NLP

Output will be the best bowler, pitch region and ground region against the batter provided as the input. Fig. 2 – 5 show the results generated using SVM, decision tree, Random Forest and NLP respectively.

IX. SOCIAL RELEVANCE

a) *Better Team Performance:* A team can enhance its performance in crucial circumstances, potentially leading to more victories, by correctly anticipating the ideal bowler to bowl to a specific batsman during the death overs. This requires careful analysis of the batsman's strengths and weaknesses, as well as the bowler's skill set and past performance in similar situations.

b) *Improved Player Performance:* Bowlers can modify their tactics and strategies to perform better against specific batsmen with the help of accurate predictions. This might lead to better performances from individual players. and ultimately improve the team's overall performance. In addition, players can use data analysis to identify their strengths and weaknesses, allowing them to focus on improving specific areas of their game. This can lead to more consistent performances and increased confidence on the field. Furthermore, the use of technology such as video analysis and wearable devices can provide players with valuable insights into their technique and fitness levels. This information can be used to make adjustments to training regimes and game strategies, leading to improved performance over time. Overall, the use of data and technology in cricket has revolutionized player performance and has become an essential tool for teams looking to gain a competitive edge in the sport.

c) *Strategic Decision Making:* Machine learning models can offer cricket coaches and captains insights and recommendations that will help them decide which bowlers to use and when. By analyzing vast amounts of data on individual players, teams, and playing conditions, these models can identify patterns and trends that are not immediately apparent to human observers. For example, they can identify which bowlers are most effective against particular types of batsmen or in certain weather conditions. They can also provide real-time feedback during matches, helping coaches and captains make strategic decisions based on the latest information. However, it is important to remember that machine learning models are only as good as the data they are trained on. Coaches and captains must ensure that the data they feed into these models is accurate and up-to-date in order to get the most value from them. Additionally, they must be willing to adapt their strategies based on the insights provided by these models, even if those insights challenge their preconceived notions or go against conventional wisdom. Ultimately, machine learning can be a powerful tool for decision making in cricket and other sports, but it must be used wisely and with careful consideration of its limitations

d) *Enhanced Fan Experience:* Machine learning forecasts may offer a new level of excitement and involvement to the game for cricket fans. They can also improve the viewing experience by allowing commentators to offer insights based on the forecasts.

X. CONCLUSION

The use of machine learning is becoming more common in practical applications. It would be feasible to analyse the performance of individual players during a tournament by adding this technology into the sport of cricket, allowing team management to make better informed judgements about which players to choose for forthcoming matches. We wanted to predict the best possible team members that could give the best possible result in a match. In the death overs, the captain of the bowling team can send their best bowlers against the respective batsman and win the match according to the way the match is being played rather than having a 50:50 chance of winning. Also, the batting team has a greater advantage when compared to the bowling team. We aim to eliminate such an advantage in the game of cricket through our project. . Our project focuses on introducing new rules and regulations that will make the game more balanced and fairer for both teams. This will ensure that both teams have an equal opportunity to win, without any undue advantage to either team. This will prevent one team from dominating the game and give the other team a chance to catch up. Overall, our project aims to make cricket a more exciting and fair game for everyone involved, while preserving its traditional charm and appeal.

REFERENCES

[1] Cricket, http://en.wikipedia.org/wiki/Cricket.

[2] Tirtho, Devopriya, Shafin Rahman, and Md Shahriar Mahbub. "Cricketer's tournament-wise performance prediction and squad selection using machine learning and multi-objective optimization." Applied Soft Computing 129 (2022): 109526.

[3] Rauf, Muhammad Arslan, et al. "Extraction of Strong and Weak Regions of Cricket Batsman through Text-Commentary Analysis." 2020 IEEE 23rd International Multitopic Conference (INMIC). IEEE, 2020.

[4] Wickramasinghe, Indika Pradeep. "Predicting the performance of batsmen in test cricket." Journal of Human Sport and Exercise 9.4 (2014): 744-751.

[5] Ahmed, Faez, Abhilash Jindal, and Kalyanmoy Deb. "Cricket team selection using evolutionary multi-objective optimization." Swarm, Evolutionary, and Memetic Computing: Second International Conference, SEMCCO 2011, Visakhapatnam, Andhra Pradesh, India, December 19-21, 2011, Proceedings, Part II 2. Springer Berlin Heidelberg, 2011.

[6] Norman, John M., and Stephen R. Clarke. "Optimal batting orders in cricket." Journal of the Operational Research Society 61.6 (2010): 980-986.

[7] Mahbub, Md Kawsher, et al. "Best Eleven Forecast for Bangladesh Cricket Team with Machine Learning Techniques." 2021 5th International Conference on Electrical Engineering and Information Communication Technology (ICEEICT). IEEE, 2021.

[8] Anik, Aminul Islam, et al. "Player's performance prediction in ODI cricket using machine learning algorithms." 2018 4th international conference on electrical engineering and information communication technology (iCEEiCT). IEEE, 2018.

[9] Ahmad, Haseeb, et al. "Prediction of rising stars in the game of cricket." IEEE Access 5 (2017): 4104-4124.

[10] Somaskandhan, Pranavan, et al. "Identifying the optimal set of attributes that impose high impact on the end results of a cricket match using machine learning." 2017 IEEE International Conference on Industrial and Information Systems (ICIIS). IEEE, 2017.