International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 06 | June - 2023 SJIF Rating: 8.176 ISSN: 2582-3930

Cricket Analysis Using Machine Learning

Harjot Kaur[1],Vaishnavi Hemade[2],Mayur Kondhare[3],Koustubh Juvekar[4],Prof. Amruta Chitari[5]

[*1, 2, 3,4]Student, [5]Assistant Professor
Department of Computer Engineering, Dr. D. Y. Patil School of Engineering, Lohegoan,
Savitribai Phule Pune University, Pune, Maharashtra, India
kaurkhushi308@gmail.com[1],
vaishnavihemade96@gmail.com[2],
mayurkondhare7875@gmail.com[3],
koustubhjuvekar2000@gmail.com[4],
amruta.chitari@dypic.in[5]

----------***----------

ABSTRACT

Nowadays we know that cricket is one the most popular and watched sport. Winning in cricket depends on various factors such as home ground advantage, performances in the past matches, weather conditions, performance at the specific venue, performance against the specific team, pitch condition. So, we have developed a system using machine learning algorithms where our work suggests the best players for the team selection and projection of score.

A subset of artificial intelligence (AI) called machine learning (ML) enables software applications to improve the accuracy of predictive values without further explanation. To predict outcomes, machine learning algorithms use historical data to generate ideas, learn, and analyze. It is a method of making predictions based on machine learning. It is a comprehensive training with all the important things like event information and history, ball loss, hometown, manager, supporters, competition, historical information history and more. They all have different strengths. Finally, our system shows the most accurate and algorithmically relevant results. In addition, it shows how well our algorithm predicts the number of goals, one of the most important aspects of a successful match.

Keywords: Cricket analysis, classification, a machine learning algorithm, and prediction.

I. INTRODUCTION

We know that machine learning is part of everyday life. The scientific study of algorithms and statistical models that computers use to perform certain tasks without being specially designed is called machine learning (ML). This short tutorial is designed to provide an overview of machine learning techniques for sports analysis, team prediction and odds. Cricket analysis systems play an important role and help solve many cricket related problems. Some of these problems include talent balance with the team, team ranking, physical and mental training preparation, developing strategies to win games and competitions, evaluating

the effectiveness of coaches and health professionals, treatment of sports injuries (health and insurance), analysis and improvisation of policies, quality of equipment and technology. , determination of prizes, history of gambling activities and establishment of odds, weather and location etc.
Using machine learning, we will develop a system that can analyze the situation, examine past performances, and predict cricket matches' scores. We are analyzing the past data and building a system which will give you best payers for the team. We have taken weather, pitch, ground etc. attributes for the analyzing process. As you know, victory in cricket depends on many important factors, such as the weakness of each opponent, the pitchers and batsmen, the overall situation of the team and the player. The most suitable and accurate algorithms in our system finally show the value proving that the team won the prediction.

Machine Learning

In this system, machine learning is used to analyze and compare players' past histories, climate, and venue characteristics. Based on these characteristics, a prediction is made about which 11 player will be the best for a particular match. When it comes to selecting better players for that particular match, this analysis is extremely helpful.

II. LITERATURE SURVEY

Sr. No.	Paper name	Advantages	Limitations
1.	Analyzing the performance of the Indian Cricket Team using Weighted Association Rule Mining	1.Accurate output 2.Easy to use interface	1. Requires high internet speed for working
2.	A Classification Based Tool to Predict the Outcome in ODI Cricket	1.Speed, Portability Efficient to use and easy interface.	1.Less Cost-efficient. 2.User needs to put correct data or else it behaves abnormally.
3.	CricAI: A Classification Based Tool to Predict the Outcome in ODI Cricket	1. Easy to use interface	1. Requires high internet speed for working
4.	Cricket Analysis using machine learning	1. SVM Algorithm	1. Not most accurate result.
5.	Review on Cricket Analysis and Prediction using Machine Learning Approach	1. Easy to use interface 2. Clustered and stored data	1. User needs to put correct data or else it behaves abnormally. 2. . Requires high internet speed for working

Table-1: Literature Survey Table

IV.SYSTEM ARCHITECTURE

The client in the system's architecture can be any user requesting a prediction of the match's score. a dataset that provides the server with the necessary player data for processing. In contrast, the user sends the information to the server, which reads it, performs the prediction and analysis, and returns the results to the user.

At the very end, a variety of machine learning algorithms are used to make predictions, with a particular prediction model being used for both the actual prediction and the data preprocessing steps. Importing the dataset, cleaning the dataset, normalizing the attributes of the dataset, and getting the dataset ready for the next processing step are all parts of the data preprocessing process.

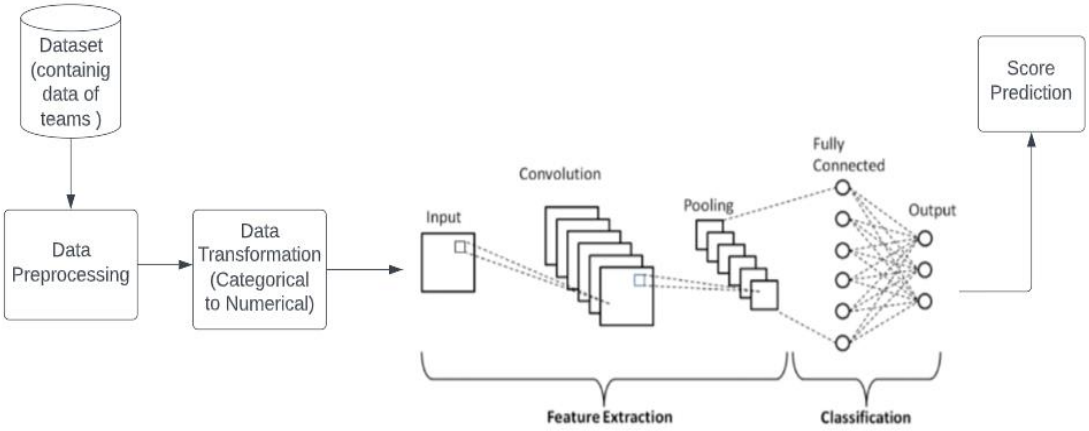


Fig.1: System Architecture

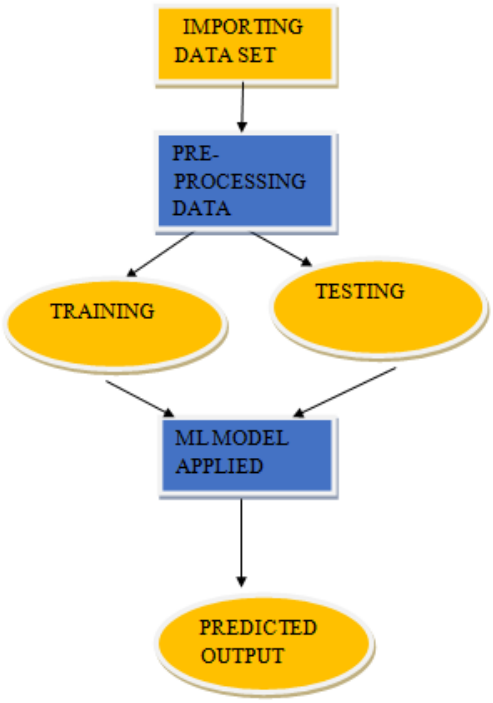


Fig. 2: Predictive system

V. UML DIAGRAMS

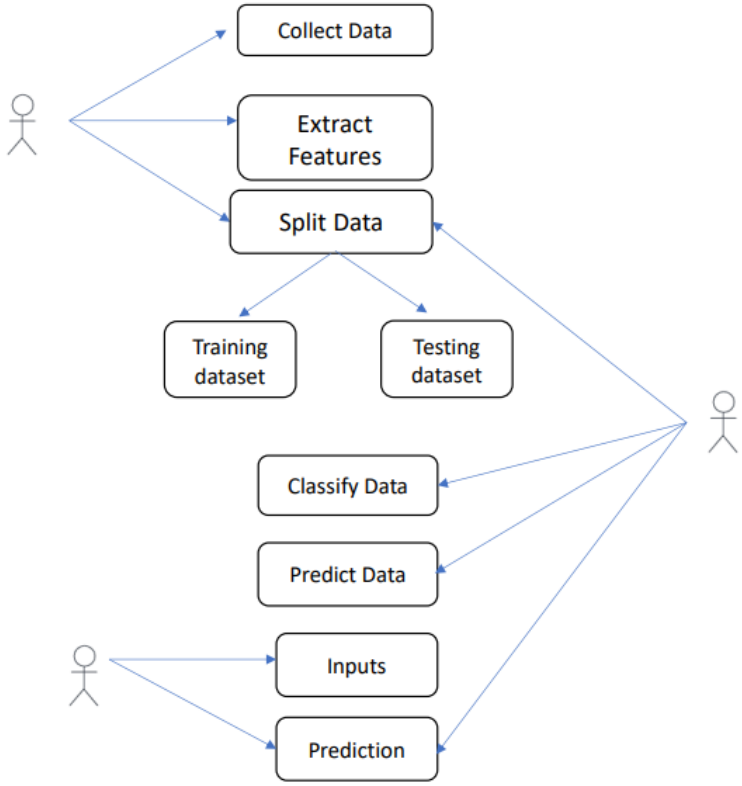


Fig. 3. Use case diagram

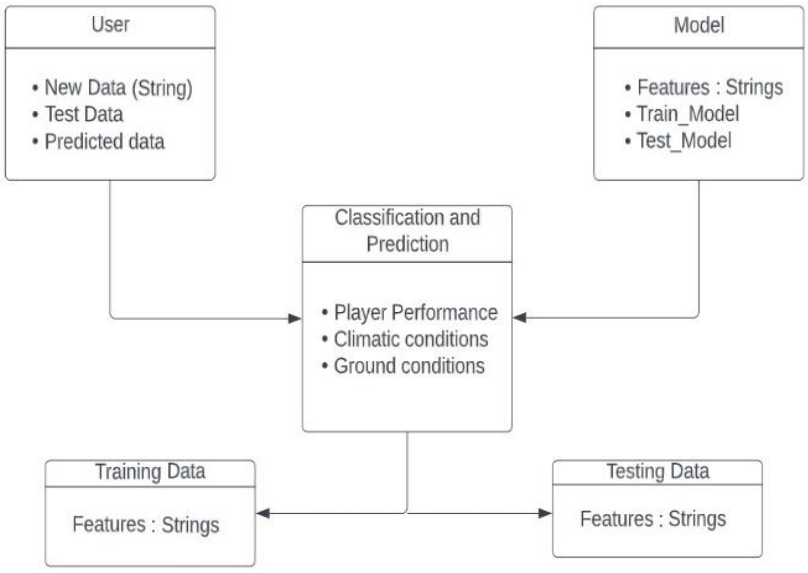


Fig. 4: Class diagram

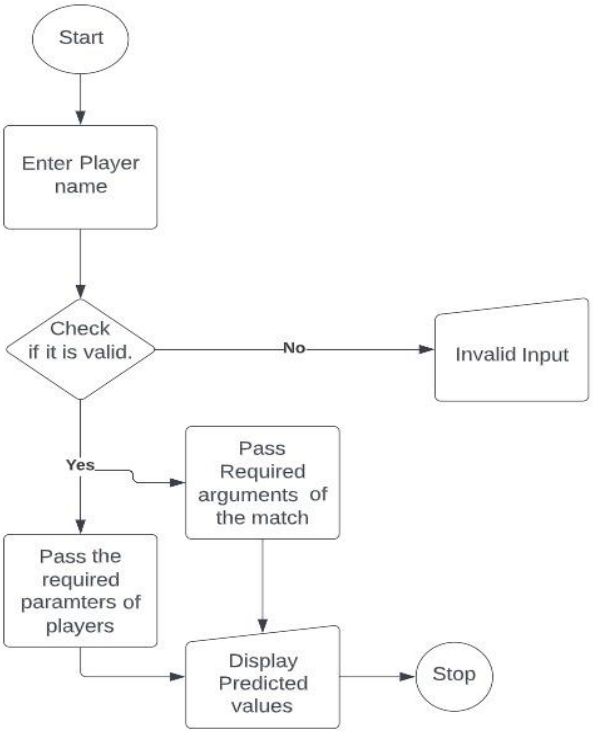


Fig. 5: Activity Diagram

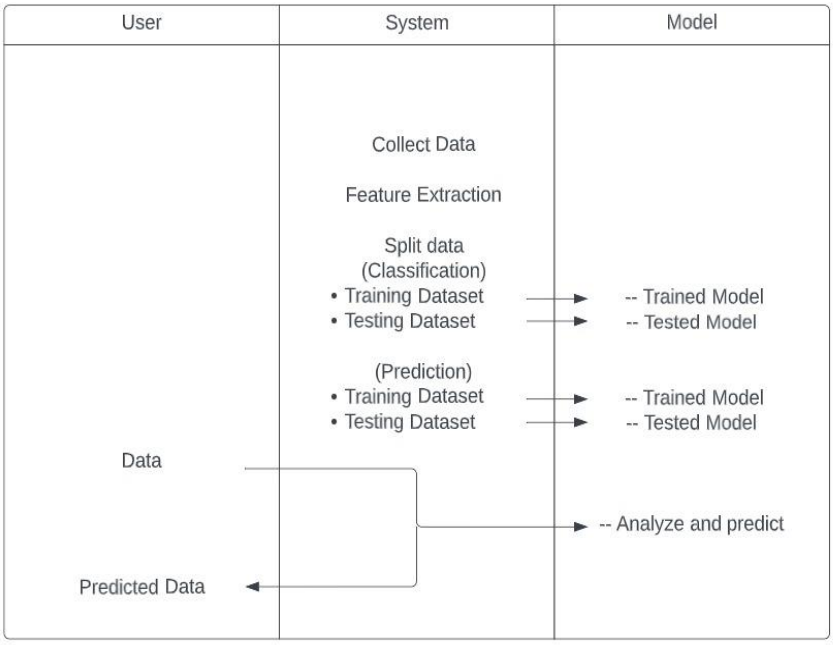


Fig. 6: Sequence diagram

VI. ALGORITHMS

Here, we discuss all the algorithm used while developing the project.

A. Decision Tree Algorithm:

Decision tree is an unsupervised learning algorithm used for classification and regression. It has a hierarchical tree structure consisting of root nodes, branches, internal and leaf nodes. The Decision Tree is an example of a graphical representation used to provide any solution to a problem/decision based on given conditions. It is called a decision tree because, just like a tree, it starts from the root node and expands into many branches to form a tree-like structure. If we are going to build a tree, we use the CART algorithm, which stands for Classification and Regression Tree Algorithm.

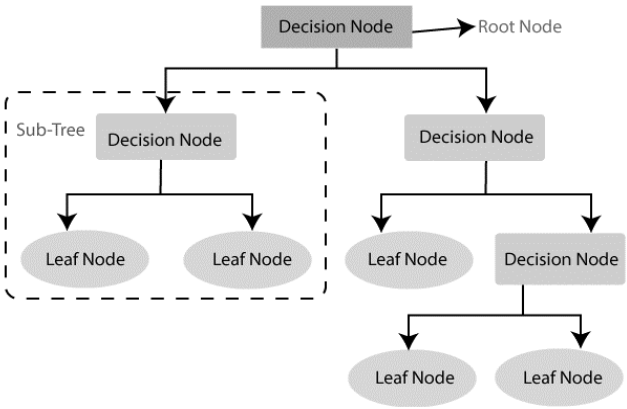


Fig. 7: Decision Tree Algorithm diagram

B. Random Forest Regression

Random forest regression is a supervised learning algorithm that uses learning techniques for regression. The unified learning process combines predictions from multiple machine learning algorithms to produce more accurate predictions from a single model.
It has many different decision trees consisting of different data and averages for accurate prediction.

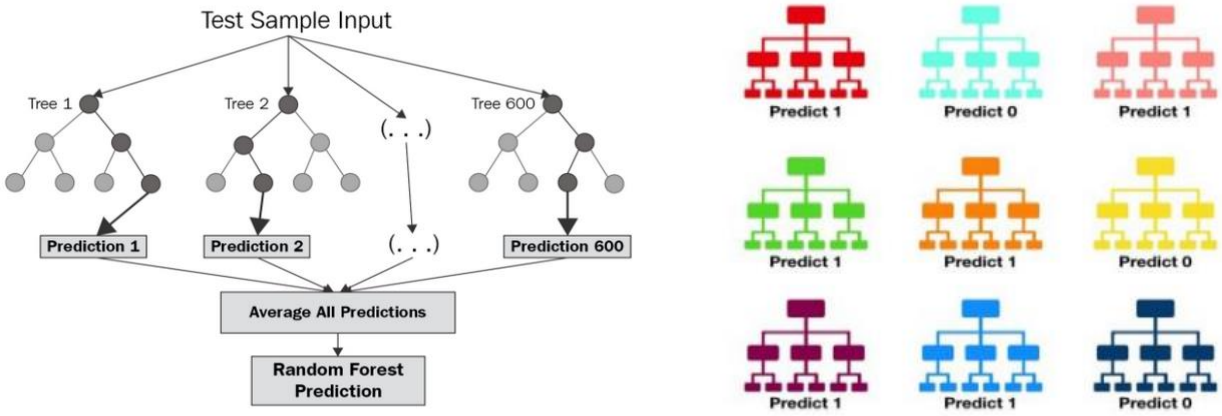


Fig. 8: Random Forest Regression

Steps Followed:

1. Randomly select k data points from the training set.
2. Construct a decision tree associated with these k data points.
3. Select the number N of trees to create and repeat steps 1 and 2.
4. For each new data point, it creates its own N-tree tree, predicts the y-value for that data point, and assigns one new data point to the average. Across all predicted y-values.

C. Linear Regression

Linear regression is a frequently used method to model the relationship between dependent variables and one or more variables. He assumed a relationship between variables and used mathematics to estimate the coefficients that fit the data.
The linear regression algorithm shows the relationship between the dependent variable (y) and one or more variables (y), hence it is called linear regression. Since linear regression shows a linear relationship between variables, it shows that the value of the variable changes as the value of the individual variable.

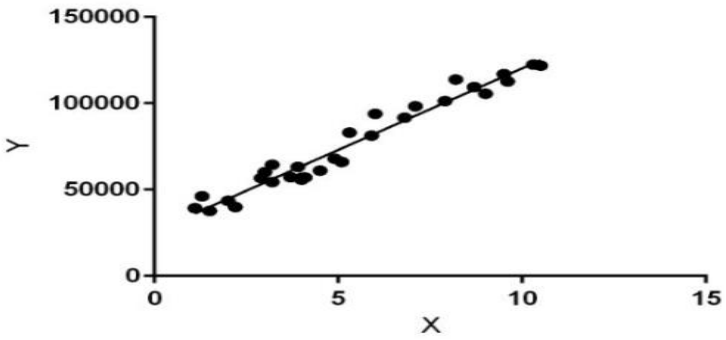


Fig. 9: Linear Regression

D. Support Vector Machine Algorithm

SVM is short for Support Vector Machine algorithm that selects point clouds/vectors that help create the plane. These conditions are called support vectors and hence the algorithm is called vector machine. SVM selects point clouds/vectors that help create a general plane. These conditions are called support vectors, so the algorithm is called a vector machine..

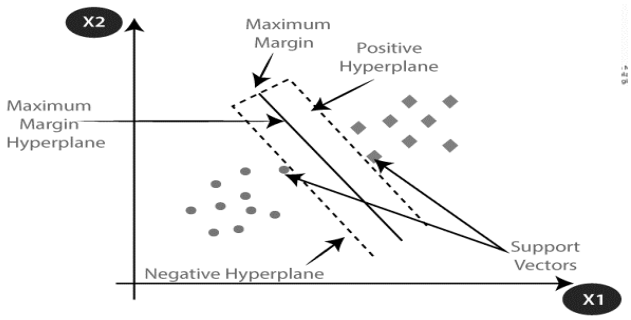
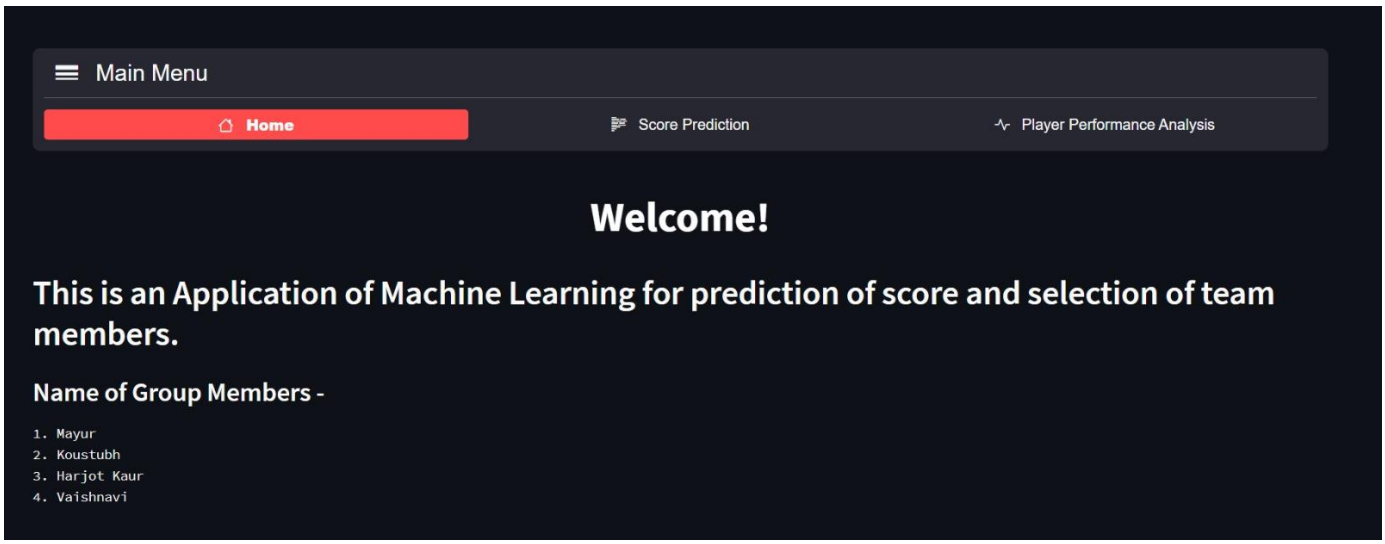
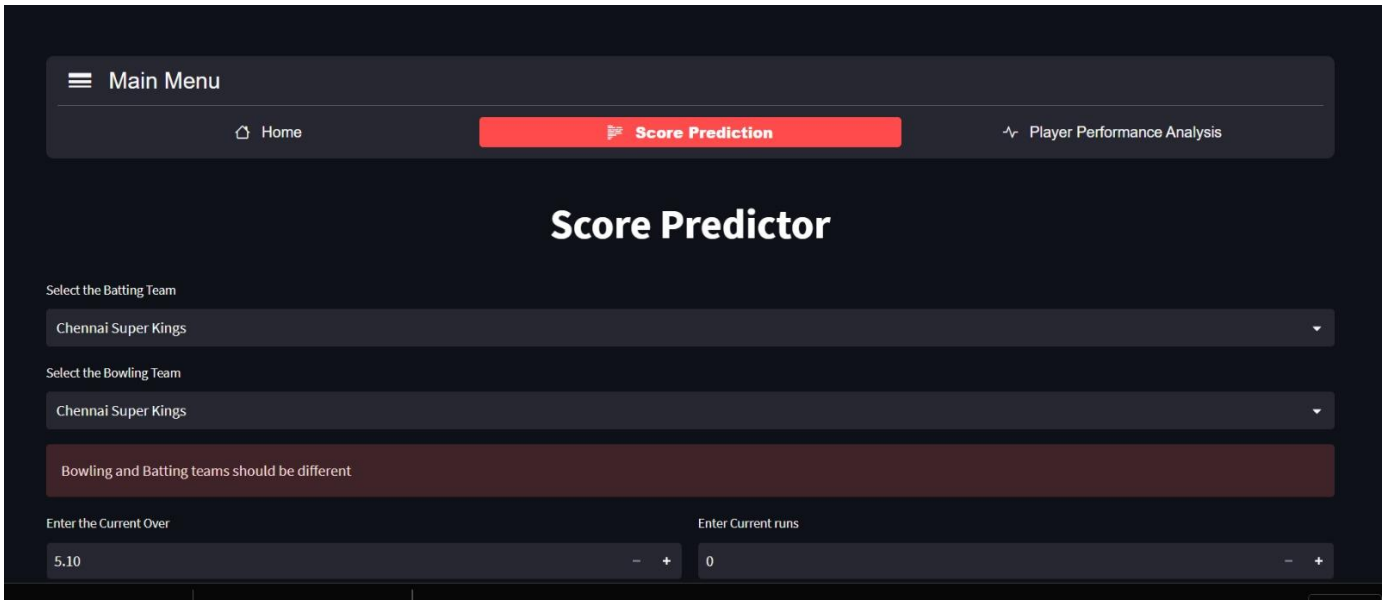
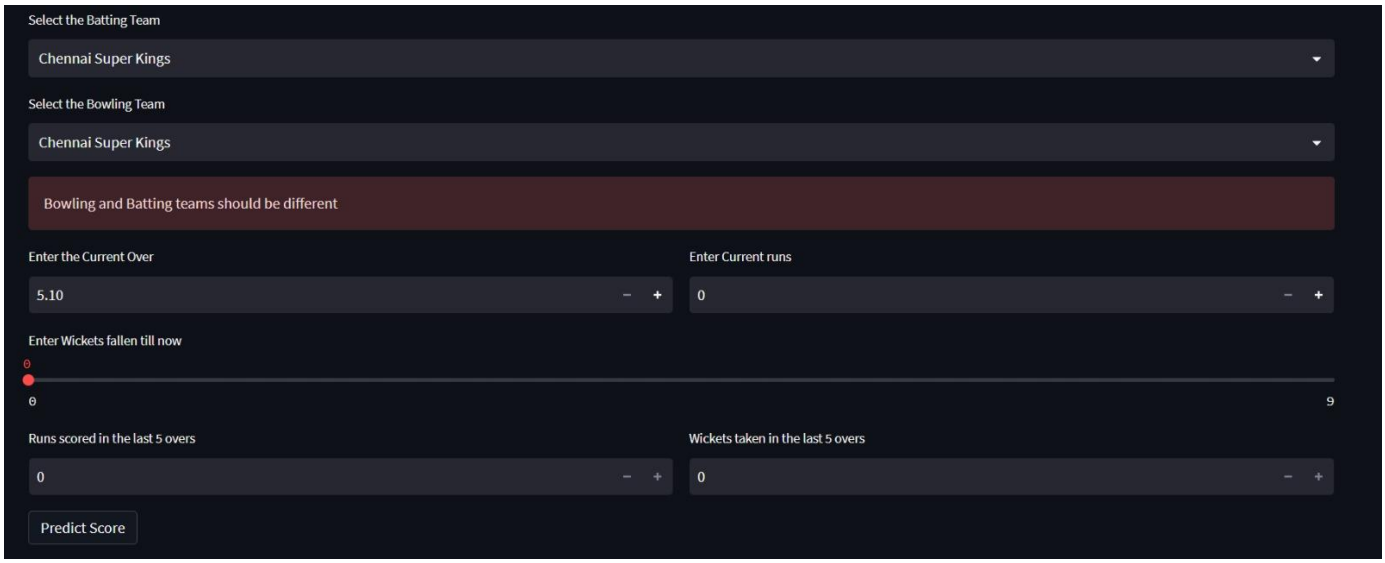


Fig. 10: SVM Tree

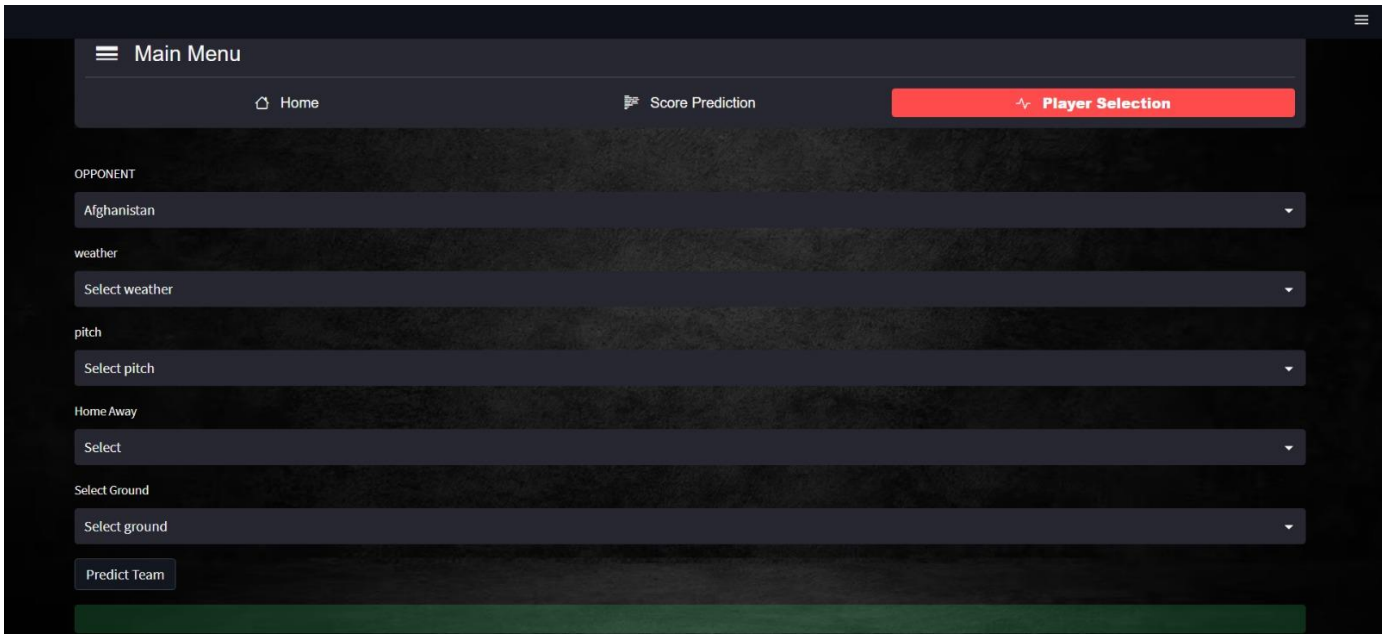
VII. IMPLEMENTATION AND RESULTS

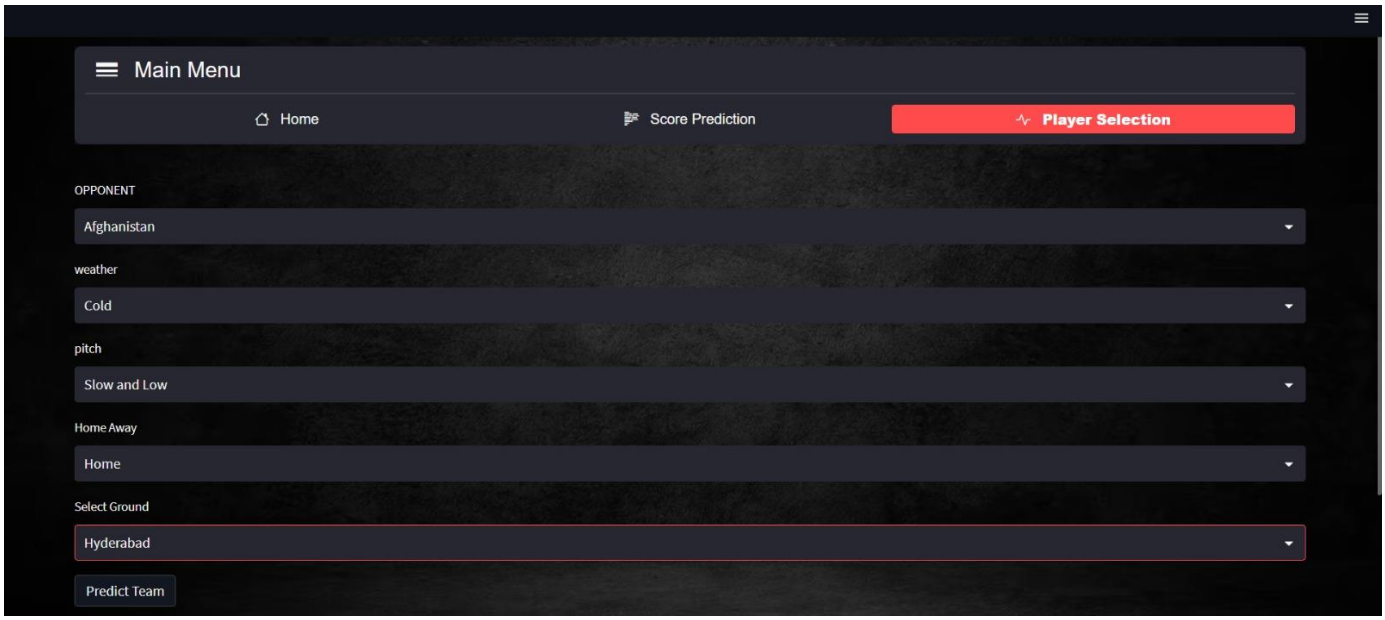


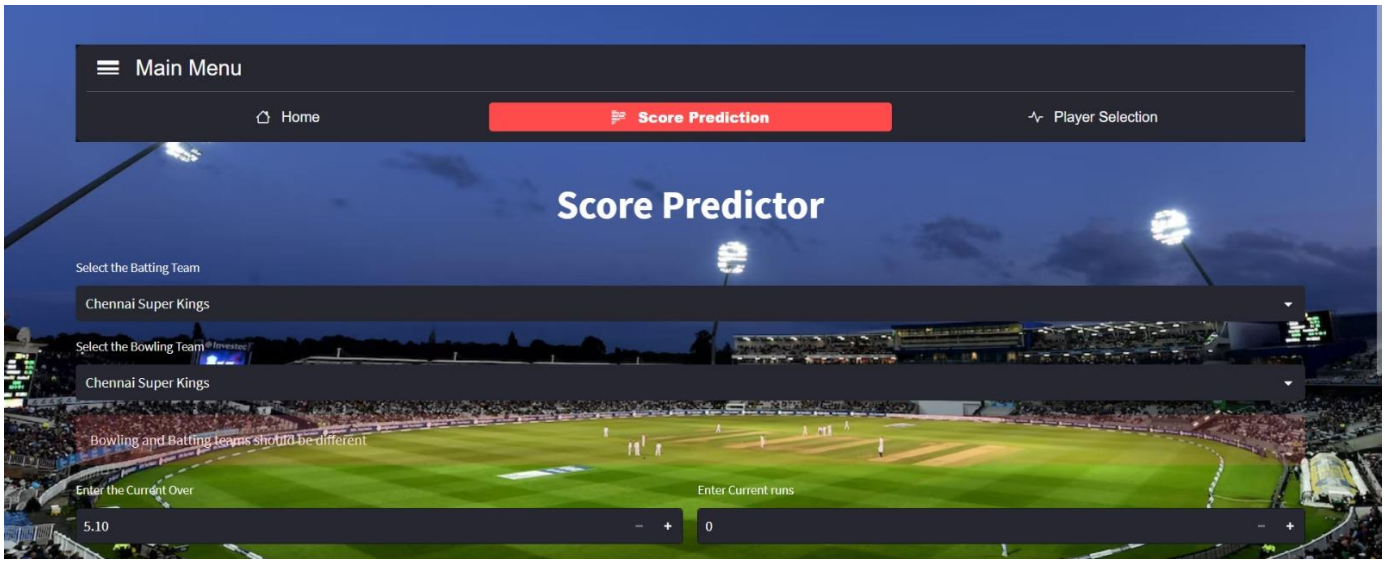


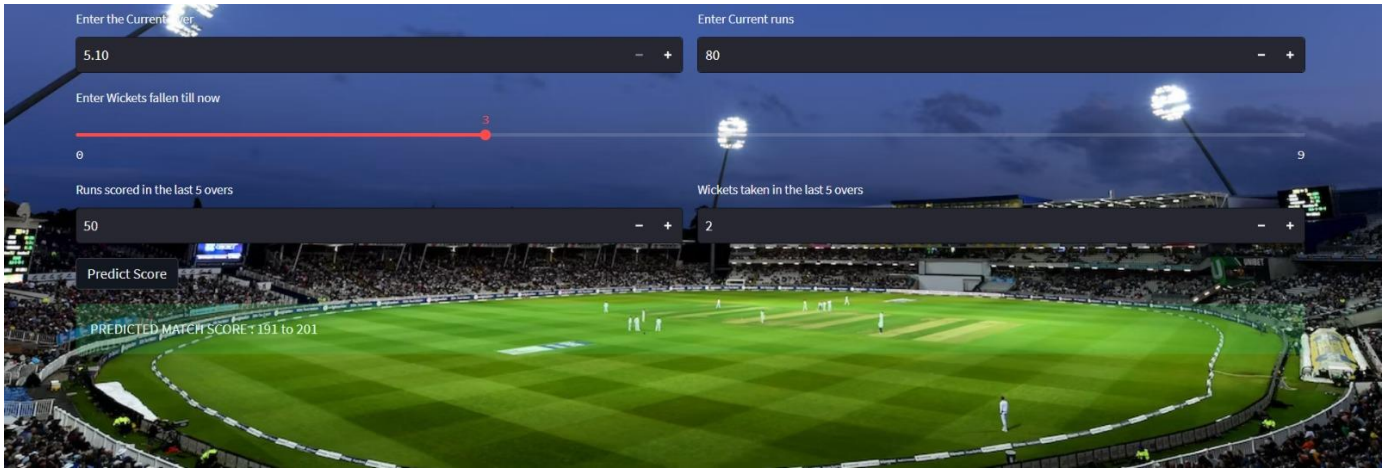

Select the Batting Team
Chennai Super Kings
Select the Bowling Team
Chennai Super Kings
Bowling and Batting teams should be different
Enter the Current Over
5.10
Enter Current runs
0
Enter Wickets fallen till now
0
0
9
Runs scored in the last 5 overs
0
Wickets taken in the last 5 overs
0
Predict Score


Runs scored in the last 5 overs
17
Wickets taken in the last 5 overs
0
Predict Score
PREDICTED MATCH SCORE : 172 to 182


Main Menu
Home
Score Prediction
Player Selection
OPPONENT
Afghanistan
weather
Select weather
pitch
Select pitch
Home Away
Select
Select Ground
Select ground
Predict Team


Main Menu
Home
Score Prediction
Player Selection
OPPONENT
Afghanistan
weather
Cold
pitch
Slow and Low
Home Away
Home
Select Ground
Hyderabad
Predict Team


Main Menu
Home
Score Prediction
Player Selection
Score Predictor
Select the Batting Team
Chennai Super Kings
Select the Bowling Team
Chennai Super Kings
Bowling and Batting teams should be different
Enter the Current Over
5.10
Enter Current runs
0


Enter the Current Over
5.10
Enter Current runs
80
Enter Wickets fallen till now
3
0
9
Runs scored in the last 5 overs
50
Wickets taken in the last 5 overs
2
Predict Score
PREDICTED MATCH SCORE : 191 to 201

VIII. CONCLUSION

We analyzed the past data and predicted scores of cricket matches with the help of Machine Learning algorithms and compared the outcome of both the methods. We also studied and analyzed some other attributes related to match and predicted best players for the match. Random Forest proved to be the most accurate classifier with the best accuracy for the data. Different tests can be done for each type of game. Models of such models can be designed to restore the main features of tennis; For example, batters want to be stronger and play longer in test matches, while T20 matches are more score and less score.
Bowlers will also have better wicket technique and more economy in Test matches, meaning they will accept fewer targets in T20 matches. In addition, attempts can be made to boost classifier accuracies and include other parameters such as location wise results, opponent team, ground conditions, player health, match fees etc.

IX. ACKNOWLEDGEMENT

It gives us great pleasure in presenting the preliminary project report on "**CRICKET ANALYSIS USING MACHINE LEARNING**"
We would like to take this opportunity to thank our Internal Guide **Prof. Amruta Chitari** for giving us all the help and guidance we needed. We are really grateful to her for her kind support. Her valuable suggestions were very helpful.
We are also grateful to **Dr. Pankaj Agarkar** , Head of Computer Engineering Department, ADYPSOE, Lohegaon, Pune and to our Project Co-ordinator **Prof. Amruta Chitari** for their indispensable support, suggestions and motivation during the entire course of the project.
We would also be grateful to our Principal **Dr. F.B.Sayyad** who encouraged us and created a healthy environment for all of us to learn in best possible way.
We also thank all the staff members of our college and technicians for their help in making this project a successful one.

X. REFERENCES

[1] P. Sri Harsha Vardhan Goud, Y. Mohana Roopa, B. Padmaja," Player Performance Analysis in Sports: with Fusion of Machine Learning and Wearable Technology" Proceedings of the Third International Conference on Computing Methodologies and Communication (ICCMC 2019) IEEE Xplore Part Number: CFP19K25-ART; ISBN: 978-1-5386-7808-4
[2] Amal Kaluarachchi, Aparna S. Varde," CricAI: A Classification Based Tool to Predict the Outcome in ODI Cricket"978-1-4244-8551-2/10/$26.00 ©2010 IEEE
[3] Deepak Saraswat, Vijai Dev, Preetvanti Singh" Analyzing the performance of the Indian Cricket Team using Weighted Association Rule Mining"2018 International Conference on Computing, Power and Communication Technologies (GUCON) Galgotias University, Greater Noida, UP, India. Sep 28-29, 2018
[4] Manuka Maduranga Hatharasinghe, Guhanathan Poravi"Data Mining and Machine Learning in Cricket Match Outcome Prediction: Missing Links"2019 5th International Conference for Convergence in Technology (I2CT) Pune, India. Mar 29-31, 2019
[5] Harshit Barot, Arya Kothari, Pramod Bide, Bhavya Ahir, Romit Kankaria," Analysis and Prediction for the Indian Premier League",2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020