Review on Cricket Analysis and Prediction using Machine Learning Approach

Patel Divya Shantilal[1], Shete Rutuja Arun[2], Diwani Jay Kantilal[3], Prof. Pawar Supriya Haridas[4]

[1]Student, Dept. Of Computer Engineering, Jaihind College of Engineering, Kuran

[2]Student, Dept. Of Computer Engineering, Jaihind College of Engineering, Kuran

[3]Student, Dept. Of Computer Engineering, Jaihind College of Engineering, Kuran

[4]Professor, Dept. Of Computer Engineering, Jaihind College of Engineering, Kuran

---***---

Abstract -*Many team sports may need a tactical assessment, movement analysis, video, and statistical database and modeling, and coach and player data presentations. Large amounts of data were collected in the world of cricket. The data requires to be analyzed for useful insights. Cricket analytic is an ideal market for acquiring innovative players . The Cricket Analysis and Prediction using Machine learning will make use of different machine learning techniques. Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. Predictions like who would be the winner, what would be a score, or which player has the same skills as other players will be made using the Machine learning model. The Proposed system will be developed using the Python programming language. Django framework will be used. The Ml learning model will be built and trained using Sklearn and Tensorflow.*

Key Words: Machine Learning, Supervised learning, Analytic, Cricket.

1. INTRODUCTION

Sports Analytics is the collection of relevant, historical, statistical data that can provide a competitive advantage to a team or individual. This data is created using sporting statistics. As far as their significance, these are mostly used in basketball and baseball. The one difference between statistical models used by analysts and advanced statistic models, however, is that the latter use probability as their main tool. Examples of statistical models include win shares, win shares of players, win shares of teams, median wins, and so on. These are statistical models that use the methods of mathematics and statistics to describe the output of a set of numbers such as votes in elections or the number of batters in a baseball lineup. These models use heuristics to provide a solution to problems, such

as the forecasting of the number of players in a lineup. Cricket analytic is an ideal market for acquiring innovative players. The teams are trying to integrate the latest tools in their infrastructure to improve the performance of their players. Cricket Analytic helps customers understand and predict the performance of cricket teams and players.

In this paper, we propose a system that will help users to find multiple types of insights from cricket data. The proposed system will make use of different machine learning techniques. Machine learning is a branch of artificial intelligence (AI) and computer science which focuses on the use of data and algorithms to imitate the way that humans learn, gradually improving its accuracy. Predictions like who would be the winner, what would be a score, or which player has the same skills as other players will be made using the Machine learning model. The Proposed system will be developed using the Python programming language. Django framework will be used. The Ml learning model will be built and trained using Sklearn and Tensorflow.

2. LITERATURE SURVEY

The Author of [1] discusses the analysis role of machine learning in the improvement of performances of players and the team in different sports and how wearable technology helps the players to know their performance levels and further improvements. The paper describes a system that collects raw data for each sport, team and player, and it is processed into statistical data. These data sets are clustered and stored as the data to be stored is very large. The paper also discusses wearable sensors. These wearable sensors are used in recognizing real-time tasks in sports. The devices are also helping the coaches in the transformation of Decision making. The paper concludes Machine Learning, along with Wearable devices can make a great impact on the players by making patterns, strategies, planning, reducing the risk of injury, and improving their performances. In [2] CricAI Toll id discussed it. It can help adjust certain factors to maximize the chances of winning the real game. The paper addressed the problem of predicting the chances of victory in a One Day International cricket match. The paper also shows a comparative evaluation of the classifiers. The systems described have been able to predict the winning criteria formulated using attributes from the dataset. The

CricAI tool can be used in real-world applications by teams playing cricket. It can help adjust certain factors to maximize the chances of winning the real game [2]. Weighted Association Rule Mining algorithm for analyzing the Indian cricket team in one-day international cricket matches against Sri Lanka and South Africa is performed and mentioned in [3]. This analysis is used by the team for framing game-winning plans. In [3] the author shows that a bigger data set can improve the accuracy of the prediction.

3. PROPOSED SYSTEM

The proposed system will make use of Machine learning algorithms.ML Algorithms will be trained using data samples. The cricket data will be scraped from the website. The scrapped data will be cleaned and stored. The cleaned data will be processed in to use for training purposes. Ml model will be trained using data. The cleaned data will also be used to generate visualization and descriptive statistics. Plotly and matplotlib libraries of python will be used to draw graphs and charts. Ml model will be implemented using Sklearn. The user of the system can set parameters for prediction. Depending on parameters and model trained the system will make predictions. The predictions will be displayed to the user. Machine Learning algorithms can be broadly classified into three categories-classification algorithms, regression algorithms, and clustering algorithms. One can leverage the advantage of all three categories in cricket analysis and prediction. The system makes use of all three categories for finding out helpful predictions. Random Forest will be used to predict the winner of the match. A lasso regression will be used to predict the score of the match and k means clustering to cluster players depending on their skills. All these machine learning models will be trained using data collected by scraping from the website. The data will be cleaned before saving to csv and using it for training. Use can check predictions by using Interface. Descriptive statistics will also be provided to the user based on historical data. The system architecture of a proposed system is shown in Fig-1below.
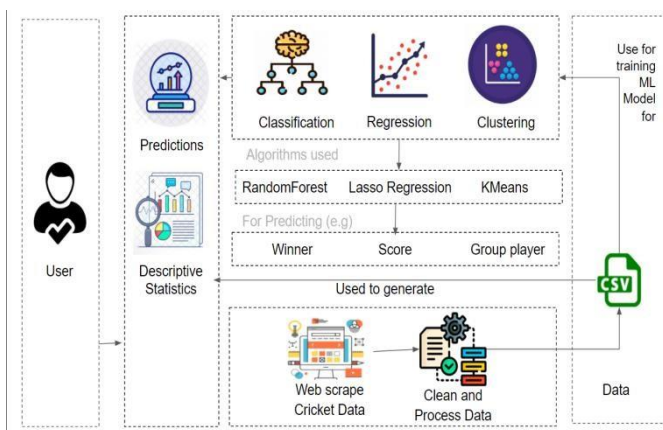
This must also be the reason why almost no attempts have been made to use Test Cricket data to make predictions, as very few Test matches are played in a year. The best format to use to model the game is T20, due to the vast number of matches played in a year. Also, T20 remains the hardest format to model due to its highly unpredictable nature compared to ODIs or Test matches. Analysis of IPL data is performed by the author in the paper [4].

In the proposed system 3 types of machine learning models will be used. A classification model will be trained to predict the winner of the match. The proposed system will use Lasso regression for predicting the score of a match, K means clustering algorithm will be used to cluster players.

4. CONCLUSION

The proposed system can prove beneficial in the cricket analytic area. It will also help to monitor the performance of a player. Clustering the same type of players will become easier.

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

Fig -1: System Architecture