International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 05 Issue: 04 | April - 2021 ISSN: 2582-3930

I-Umpire prediction, analysis and decision making review system based on their domain using twoData structure- decision tree and linear regression

Manan Patel, Deepesh Naik, Ayush Chitnis and Mrs. Vaishali kosamkar
Department of Computer Engineering
Shah and anchor kutchhi engineering college, Mumbai University,
India

***Abstract*—It is python OpenCV, Machine Learning based application that is used to take difficult on field decisions by altering the video collected by web cam. It then adds the decisions made to a .txt file. This file contains various parameters which can be used to predict the later results. For eg: the I-umpire used by a bowlers perspective can be useful to keep tracks of no-balls, wide-balls and no of boundaries given by the bowler and thus can be used to predict the type of ball he will bowl in the next match. The main objective of this project is to explore different Machine Learning techniques to predict the score and outcome of cricket matches, using in-game match events rather than the number of goals scored by each team. We will explore different model design hypotheses and assess our models' performance against benchmark techniques. In this project, we developed an 'expected wides' metric which helps us evaluate a team's performance, instead of using the actual number of runs scored. We combined this metric with a calculation of a team's offensive and defensive ratings which are updated after each game and used to build a classification model predicting the outcome of future matches, as well as a regression model predicting the score of future games. Our models' performance compare favourably to existing traditional techniques and achieve a similar accuracy to bookmakers' models. Our proposed system grades the prediction in the topic "Data Structures" with a domain accuracy of 89.47 percent and language accuracy of 84.79 percent.**

***Index Terms*—domain accuracy, language accuracy, prediction and analysis, data structures.**

I. INTRODUCTION

The introduction and generation of interesting new ideologies from the human brain has reached an all-time high thanks to the transcending world of technology emerging from Internet Technology. Furthermore, when looking at it from a global perspective, there have been some fascinating trends, as shown by the major fields of consumer goods, sports etc.There are countless possibilities for how the internet connects the dots, regardless of the type of any sport, ranging from football to basketball and basketball to baseball. This infiltration among sports has led to its own distinct presence, which could be referred to as the "Internet of Sports." One of the biggest mutation in "Internet of Sports" has not only put the financial success stories in the codex of various professional athletes but also evolved the company on the path to grandiose triumph. Activities and subjects of interest like Statistical calculation of the athletes was seldom worked in an organized and sequential manner back in the ages where these concerns were age old conundrum. With the help of the power of the Internet, the everyday analysis of these dilemmas seemed so easy to implement and simple to observe. .Some other acknowledged indistinguishable transcendence in the world of the Internet have also ranked at the status where the work has been tirelessly done and serving as a stepping stone for success in terms of revolution with the period, and obviously the potential has been seen when it comes to sportsThe rest of the paper is divided as follows. Section 2 gives an insight of the existing literature and the research gaps. Section 3 discusses our approach to bridge the research gaps. Section 4 describes the detailed

implementation of our system and results are provided in section 5. Following which, section 6 analyzes the results before drawing the final conclusion in section 7.

II. RELATED WORK

The recent survey paper(s) to predict the outcome of English country twenty over cricket matches[1],[5] Forecasting techniques were implemented to describe the studied event of FIFA World Cup 2006 and prediction markets were introduced for prediction of results of all the matches held in the tournament. For prediction market, they created a web interface that ran in coherence with the time the tournament took place. They calculated some virtual 7 currency units that had certain values for every match win or loss. The trading screen interface was well evaluatedResearchers used some datasets as a new benchmark, which happened after forecasting on historical data that has been done by them. The second benchmark was on the basis of ranking of teams by FIFA while considering random predictor as a primary one. For accuracy comparison, the percentage of correctly predicted games were reckoned by the researchers and this data was stored as hit rate (33.3% for random draws). Greater the hit data, more are the chances for a team to win a match.

For research methodologies, researcher deduced an overall plan to get the resultant decisions of people. Plans included various descriptive studies, with longitudinal research designs. Using descriptive research design, the trend was projected of sports betting growth and its effect on distinct groups[2].Other methodologies used were quantitative approach, to quantify data in numerical quantities. These methodologies helped researchers to study the statistical measurable variables. 8 Sampling design categories (Frame and Technique) were referred and were used to determine the sample of population, its selection of observable sample and the tolerance of margin of error in selected sample. This study used various formulas to come up with adequate sample size. Furthermore, some other data collection methods were used as an amplification factor for the data collection.

Muthuswamy and Lam[1] predicted the performance of Indian bowlers against seven international teams against which the Indian cricket team plays most frequently. They used back International Journal of Data Mining & Knowledge Management Process (IJDKP) Vol.8, No.2, March 2018 20 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. Wikramasinghe[2] predicted the performance of batsmen in a test series using a hierarchical linear model. Barr and Kantor[3] defined a criterion for comparing and selecting batsmen in limited overs cricket. They defined a new measure P(out) i.e. probability of getting out and used a twodimensional graphical representation with Strike Rate on one axis and P(out) on another. Then they define a selection criterion based on P(out), strike rate and batting average of the batsmen. 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 the next match.

Recommender Systems (RS) are helping internet users as expert. They help the users in makingthe decision in day to day life about what to purchase, which show to watch, which place to visit etc. Recommender systems (RS) try to identify the needand preferences of users, filter the huge collection of data accordingly and present the best-suited option before the users by using some well-definedmechanism (Sohail, Siddiqui, & Ali, 2017). RS are beneficial to both the users and service providers (Pu, Chen, & Hu, 2011). They have also proved theirimportance in decision making and have improved the quality of decision making. Neural network technique was used to develop a model for player selection, in this model different attributes needed for player selection is analyzed

into four major categories using a neural network model and this major categories which include the player's technique, the player's speed, the player's physical status and the player's resistance

III. PROPOSED METHODOLOGY

In this section we describe the proposed method for prediction and analysis model subjective answers on the topic data structures. The system with the help of machine learning modules like decision tree and linear regression helps to assign the prediction and analysis module. The aim of the system is to provide the proper results by which the selection of players can be done.

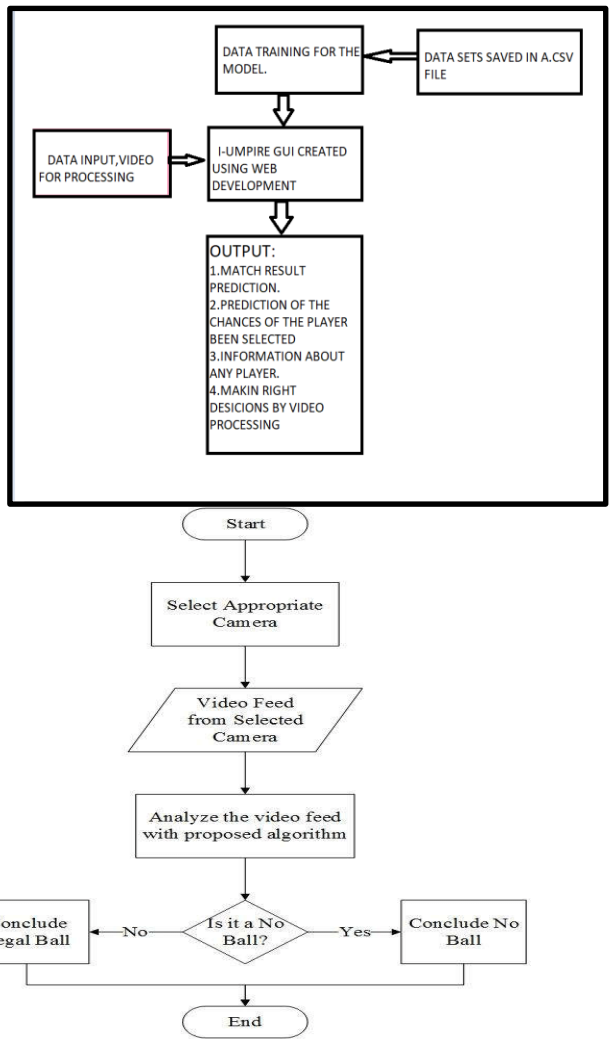

Fig.1The flow of output using the data input and training model for accurate results giving the input to an output tuple.

A. *Input Block*

The input block which is the input stage where all the data sets of our project is been saved in csv format file. The datasets are available in excel format and taken as input for the results as prediction and for analysis. The next stage which is data training model in which data is taken in and further trained for resulted output for much consistent and accurate results for our project.

B. *data input and video processing Model Block*

The data input from the prepared datasets or the recorded video which is saved in our device or a live recorded video which can be used for i-umpire GUI created model which can be quite helpful to provide much accurate result.

Fig.2Example for no ball decision

C. *I-umpire GUI model*

In this module, the data input and video input can be taken from the data stored or video can also be take live for the GUI model in which it processes the video and the following video can be resumed , fast forward , slow or out- not out decision can be taken easily and can be used to take tough on-field decision and for accurate result.

D. *Output Block*

Based on the datasets and the input video, the output model processes the input and following output results can be obtained which are accurate

and most predicted results. The output can provide us match prediction result, bowler or batsmen analysis with the help of it we can decide whether a player can be selected or not and the decision review system which help us to provide the tough on-field calls/decision easily and with great accuracy.

IV. IMPLEMENTATION

The implementation started with searching for the necessary datasets that would satisfy our model. We found one file from the Kaggle website about the IPL matches played all over the years.This file also contains various parameters which can be used to predict the later results. For eg: the I-umpire used by a bowlers perspective can be useful to keep tracks of no-balls, wide-balls and no of boundaries given by the bowler and thus can be used to predict the type of ball he will bowl in the next match. We found a detail Dataset about all the matches played from 2008 to 2017 along with the players information and information about weather and pitch report and on the basis of that we were able to classify players like batsmen and bowlers and used them for our players selection respectively. For Match prediction we used five parameters like where the match is played the home team and opposition team who won the toss and what did the team decided to do. And on the basis of these we generated scores Later on we match that scores with another scores generated by pre solved data like No of foreign players in each team which team has more Batsmen as compared to bowlers and various other parameter and thus on the basis of that scores we successfully predicted which team has a higher chance of wining. In this project, we also developed an 'expected wides' metric which helps us evaluate a team's performance, instead of using the actual number of runs scored. We combined this metric with a calculation of a team's offensive and defensive ratings which are updated after each game and used to build a classification model predicting the outcome of future matches, as well as a regression model predicting the score of future games. Our models' performance compare favourably to existing traditional techniques and achieve a similar accuracy to bookmakers' models. Our proposed system grades the prediction in the topic "Data Structures" with a domain accuracy of 89.47 percent and language accuracy of 84.79 percent. Thus with the help of Decision Tree Classifier and Linear Regression Algorithm we were able to predict the Match results and can do the proper selection of player.

V. RESULTS

The model can be used to predict the Match Result between any two selected Teams. The Result depended upon the values selected by the user and the parameters set by the Model. Thus it shows which teams has more chances of winning on the basis of the selected values. It also provides the information about any player you want. It gives the probability of a Batsmen and a Bowler to play the next match according to the values given by the user. The VideoHelp feature helps in giving the onfield decision by Video Rendering so that the user can fast forward or slow down the video and come to some conclusions.

VI. DISCUSSION

Thus the I-Umpire can be helpful with providing Match Prediction , Player Selection, Information about any player played in the tournament and to carry out difficult on field decisions. Thus I-umpire strives to bring all the players together and exposing them to the sports. It can also help the individual to develop there own skillsets and to know there weakness and strength so that they can improve themselves. I-umpire can further be improved and expanded to different fields like different sports like Football, basketball etc. It can also be expanded to different leagues all over the world. Apart from sports it can also be used in making Models such as Weather Prediction, Stock Market Analysis etc.

VII. CONCLUSION

In this paper, we have proposed an innovative use of data structure like decision tree module used to predict the match outcome and linear regression which is used for bowling and batsman analysis purpose using which we can choose whether to select a player or not with the help of datasets and get accurate results for the following. The decision making review system developed can be used to call on tough on ground decision and would be quite helpful for the following and using the datasets prepared can be used to provide accurate and consistent decision and results for the same. Our knowledge referring to various papers helped us a lot to developed such a model and even for the research paper.Match prediction whether its winning or losing it's a very interesting topic in the field of machine learning .Selection of the right players for each match plays a significant role in a team's victory. An accurate prediction of how many runs a batsman is likely to score and how many wickets a bowler is likely to take in a match will help the team management select best players for each match. In this paper, we prepared bowling datasets based on players' stats .With the increase in number of competitions being held and the increasing need for delivering correct and swift decisions I-umpire would be a substitute to deliver these decisions and unsure the game is carried out in a fair manner. Human led errors can be avoided making the game fair for all. The future work that could be done on the system is that instead of only using a single sport we can embedthe prediction module for many other sports and can add various features for live screening and merchandise store for sport equipment too.

REFERENCES

[1] Stylianos Kampakis,"using machine learning to predict the outcome of
english country twenty over cricket matches", 2015.

[2] Robin Praet , "predicting sports result by using recommendation
techniques",2016.

[3] L. Breiman, [Random Forests], "machine learning",Oct 2001.

[4] O. Chapelle and Y.Chang, "journal to machine learning research", 2011

[5] Chao-Ying Joanne Peng, Kuk Lida Lee, Gary M. Ingersoll, "an
introduction to logistic regression analysis and reporting", September
2002.

[6] D. Parker, P. Burns and H. Natarajan, "player valuations in the indian
premier league",Frontier Economics, vol. 116, October 2008.

[7] C. D. Prakash, C. Patvardhan and C. V. Lakshmi, "data analytics based
deep mayo predictor for ipl9", International Journal of Computer
Applications, vol. 152, no. 6, pp. 6-10, October 2016.

[8] M. Ovens and B. Bukiet, "A Mathematical Modelling Approach to One-
Day Cricket Batting Orders",Journal of Sports Science and Medicine, vol. 5,
pp. 495-502, 15 December 2006.

[9] 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.

[10] D. Parker, P. Burns and H. Natarajan, "Player valuations in the
Indian Premier League," Frontier
Economics, vol. 116, October 2008.

[11] C. D. Prakash, C. Patvardhan and C. V. Lakshmi, "Data Analytics based
Deep Mayo Predictor for IPL-9," International Journal of Computer
Applications, vol. 152, no. 6, pp. 6-10, October 2016.