T20 Cricket World cup winning prediction

Agrajeet Vaze[1],Swarup Patil[2] , Dr.Manisha Bharati[3]
[1,2,3]*Department of Technology, Savitribai Phule Pune University, India*

Abstract *Cricket is a game where we need field, player, coach,support team, skills and team work. Where we can analyse, The overall perspective behind Cricket (T20) analytics is using Machine Learning (ML), Support Vector Machines (SVM), Time Series Analysis (TSA) and scope of cricket in over the world. Historically there are empirical decisions and the uses of analytics has been using intensely. From the datasets we access features such as performance of the players, individual skills and team work done by players.*
The strategy behind T20 analysis elaborates the contribution towards success rate of predicting the various possibilities and team strategies to end up with great win in that particular match against particular team. Here we are with future prediction which team has abilities to win this 2022 T20 world cup which is going to held in Australia.

Key Word*: Cricket,T20, Machine Learning, Support Vector Machines, Time Series Analysis*

I. INTRODUCTION

When the Benson & Hedges Cup ended in 2002, the ECB needed another one day competition to fill its place. Cricketing authorities were looking to boost the game's popularity with the younger generation in response to dwindling crowds and reduced sponsorship. It was intended to deliver fast-paced, exciting cricket accessible to thousands of fans who were put off by the longer versions of the game. Stuart Robertson, the marketing manager of the ECB, proposed a 20 over per innings game to county chairmen in 2001 and they voted 11–7 in favour of adopting the new format The first official Twenty20 matches were played on 13 June 2003 between the English counties in the Twenty20 Cup. The first season of Twenty20 in England was a relative success, with the Surrey Lions defeating the Warwickshire Bears by 9 wickets in the final to claim the title. The first Twenty20 match held at Lord's, on 15 July 2004 between Middlesex and Surrey, attracted a crowd of 27,509, the largest attendance for any county cricket game at the ground other than a one-day final since 1953.

Soon after with the adoption of Twenty20 matches by other cricket boards, the popularity of the format grew with unexpected crowd attendance, new domestic tournaments such as Pakistan's Faysal Bank T20 Cup and Stanford 20/20 tournament, and the financial incentive in the format.

The West Indies regional teams competed in what was named the Stanford 20/20 tournament. The event was financially backed by convicted fraudster Allen Stanford, who gave at least US$28,000,000 funding money, the fruit of his massive Ponzi scheme. It was intended that the tournament would be an annual event. Guyana won the inaugural event, defeating Trinidad and Tobago by 5 wickets, securing US$1,000,000 in prize money. A spin-off tournament, the Stanford Super Series, was held in October 2008 between Middlesex and Trinidad and Tobago, the respective winners of the English and Caribbean Twenty20 competitions, and a Stanford Superstars team formed from West Indies domestic players; Trinidad and Tobago won the competition, securing US$280,000 prize money. On 1 November, the Stanford Superstars played England in what was expected to be the first of five fixtures in as many years with the winner claiming a US$20,000,000 in each match.

In recent years, because of work and research in reality, most focus has been paid to the research and application of machine learning at home and abroad. The Machine Learning is applied to the prediction of the cricket T20 championship, and by forecasting or predicting from historical data we are making the data garrulous.Like the previous T20 World Cup 2021, also the up-coming tournament in Australia has caught the attention of several modelers who try to predict the tournament winner. They had forecasted the complete tournament by giving result such as Australia to win the T20 World Cup 2021

In this paper we compare both rating systems in terms of their historical performance (team records, batting records, bowling, Wicketkeeping records, Fielding, Partnership records, Individual records(Captains,Players,Umpires) on previous T20 Worldcup. We forecasting the wining team for the T20 World Cup 2022 Australia.

II. MATERIAL AND METHODS

Predictive Sports Analysis
Predictive Analytics is an uprising concept, which uses various techniques such as Data Mining, Machine Learning and Modelling, to train itself from the past records, and analyze them to predict the future results. Data Analytics consists of three types of models Descriptive, Predictive and Decision Models. Descriptive models are applicable to the problems of

classifications, where certain data requires grouping into certain clusters. Decision Models makes relations between the decisional elements – Known Data, Decision and Forecast Results. Our case is an example of the Predictive Model, which relates one or more known attributes of the unit and the performance of a unit in a sample.

Deep Learning for Sports Prediction
Deep learning is a type of machine learning, which is usually implemented using neural network. It is based on multi-layer feed forward networks initiated with input layers for matching the features of the dataset, followed by various layers of non-linearity and terminating with linear regression or layer of classification to map with outcome related space. The basic framework of multi-layer neural networks supports in accomplishing deep learning tasks. Accurate sports predictions, based on high level of non-linear outcomes, which require deep learning architecture-based models comprised of hierarchical feature extraction. The ability of Deep Learning model is to acquire needful insights in our dataset its T20 world cup statistics.

Literature Review
The literature on modeling the outcomes of cricket games can be divided in two broad categories, such as score-based, bowling-based, age-based and result-based models. The initial observation in models shows the number of score scored and agree by both competing teams.Although various types of machine learning models are most common in the literature ,so machine learning model could be used. For example. Groll et al. [7] found that a random forests model generally outperforms the conventional regression methods. Another popular class of models are Bayesian network. While the above learning focus on the actual prediction models, other learning have investigated the state of possible predictor. Bookmakers' odds are a first popular covariate. They reflect the predictions of bookmakers [8], who have strong economic incentives to make accurate predictions.
The cricket player's technical skills are directly concern with information and skill, proposes that various visualization tools used in Business Intelligence & Analytics techniques can represent about athletes good in their pursuit of positive achievements. Also, individual athletes are eager to leverage their athletic performance in their respective sport and aspire to be good at what their sport demands in terms of physical and competitive accomplishments..Also,(icc) focuses on the basic task in sports activity and performance analysis is to understand the relationship between development of skill and the development of players to achieve sports excellence. Hence, it is essential to develop theoretical principles to guide the concession of skill acquisition programs. The improvements for decision-making and regulation of action in dynamic environments, for instance in football, come out from the continuous performer-environment interactions.
In this method we split the data as training and testing model in the ratio 0.8:0.2 or 0.7:0.3. The only drawback of these model that if the training data includes a given

Dataset & Methodology
In this section, we briefly describe several different methods that generally come into consideration when the runs scored in single matches are directly modelled. Actually, all of them have already been used in former research on cricket data and, generally, all yielded satisfactory results. However, we aim to choose the approach that has the best performance regarding prediction and then use it to predict the T20World Cup 2022.
**Naïve-Bayes:**
Naïve-Bayes are supervised learning algorithm based on applying bayes theorem with the naïve assumption of attribute value which never occurs in the whole dataset hence the conditional probability automatically tends to zero, the model returns with biased values.

III.RESULT

Overall T20 World Cup we come to know that only few countries are getting qualified for semi-finals and as we observed that these teams are playing semi-finals for last 3 to 4 T20 World Cups.
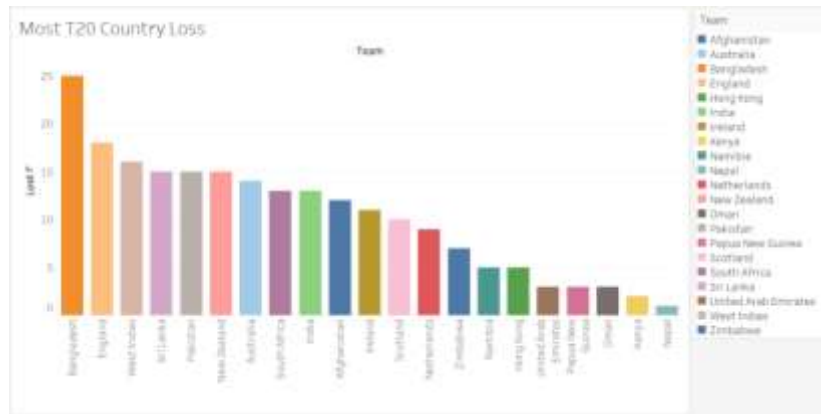
We also come to know that the countries that hosting the T20 World Cup are getting some beneficial surrounding for winning matches for their countries as per the previous results. The wining probability of countries who hosting the T20 World Cup till now is getting 30% percent chances more as compare to other component teams.
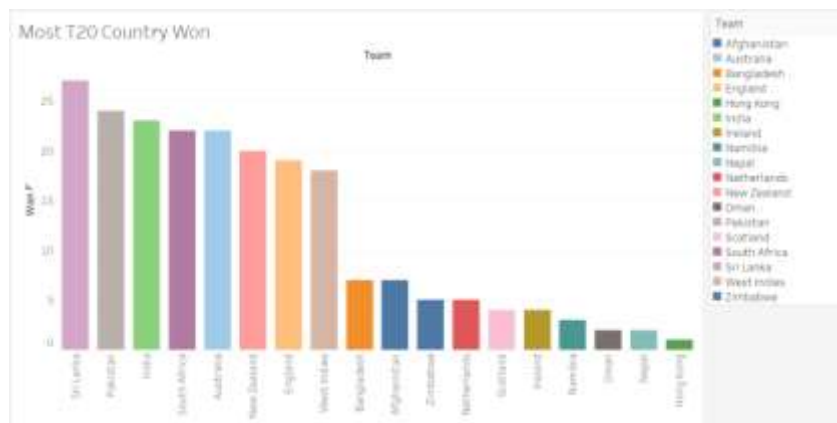
IV.CONCLUSION

The upcoming T20 World Cup is going to held in Australia. So we conclude that the countries who really getting the beneficial surroundings they probably wining the matches for there countries. Such as, as we know that , Australia, New zealand, England had got the beneficial surrounding as compared to other countries. As we compared in visualization in our Tableau tool it come to know that India, Pakistan, Australia, South Africa and Sri Lanka are the upcoming best teams for the T20 World Cup 2022.

References

[1] Bailey, M.J. & Clarke, S.R.: Market inefficiencies in player head to head betting on the 2003 cricket world cup. In Economics, Management and Optimization in Sport, S.Butenko, J.Gil-Lafuente & P.M.Pardalos, editors, Spinger- Verlag, Heidelberg,pp. 185-202 (2004).
 [2] Bailey, M.J. & Clarke, S.R.: Predicting the match outcome in one day international cricket matches, while the match is in progress. Journal of Science and Sports Medicine, 5, 480-487(2006).
[3] Barr, G.D.I. and Kantor, B.S.: A criterion for comparing and selecting batsmen in limited overs cricket, Journal of the Operational research society,55,p 1266-1274(2004)
[4] C. M Bishop, "Pattern Recognition and Machine Learning", Springer New York, 2006.
[5]RandomForest,www.stat.berkeley.edu/~breiman/RandomForests/cc_home.htm
[6]BalancedAccuracy,http://onghome.my/papers/brodersen10post-balacc.pdf
[7]https://cran.rproject.org/web/packages/gWidgets/index.html
[8] Wikipedia on the Game of Cricket, website [Online] http://en.wikipedia