IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

Analysis and Predictions of Winning Indian Premier League match using Machine Learning Algorithm

Mr. Om S. Nagare[1], Mr. Yash B. Kadve, Mr. Shivam L.Borade[3], Mr. Rushikesh N. Patil[4]
Prof. M. P. Deshmukh[5]
Department of Information Technology[1,2,3,4,5]
Matoshri Aasarabai Polytechnic, Eklahare, Nashik, Maharashtra, India

Abstract: *India's most popular sport is cricket and is played across all over the nation in different formats like T20, ODI, and Test. The Indian Premier League (IPL) is a national cricket match where players are drawn from regional teams of India, National Team and also from international team. Many factors like live streaming, radio, TV broadcast made this league as popular among cricket fans. It is an model developed to predict the winner of an ipl match using the previous data set . In this model we have used the technologies like prediction algorithm , analyzing algorithm in python language on the basis of previous 10 years ipl match data set which is collected by kaggle.com basically it contains details like team played , venue , winner , toss winner , toss decision win by runs , win by wicket and many other about 18 different fields . at the end the model gives us the accuracy of about 94 %.*

Keywords: Data analysis, prediction model, IPL Data analysis, Visualization of data, processing technique, Model Evaluation & development , mechanic learning prediction, Dateset manipulation, Accuracy, match winner, IPL team, player Performance ,previous performance, Machine Learning, Random Forest, Classification Algorithm, Cricket Prediction

I. INTRODUCTION

Indian Premier League (IPL) is a national cricket match where players are drawn from regional teams of India, National Team and also from international team. It is based on 20-20 format and is owned by Celebrities, Businessmen and others and the entire IPL is controlled by Board of Control for Cricket in India (BCCI). For the current year (2021) there are total of 8 Teams in IPL namely, Royal Challengers Bangalore (RCB), Rajasthan Royals (RR), Chennai Super Kings (CSK), Mumbai Indians (MI), Kolkata Knight Riders (KKR), Delhi Capitals (DC), Punjab Kings (PK) and SunRisers Hyderabad (SRH)

We are trying to find out the match winner of an IPL match based on the stadium they are choosing and the toss decision using machine learning techniques like Random Forest, Logistic Regression etc. Remainder of the paper is organized as follows: The section 2 is the literature survey, section 3 deals with the problem definition and the architecture. Section 4 deals with the experimental results. Section 5 talks about the conclusion.

II. LITERATURE SURVEY

Rameshwari Lokhande et al [01] the project aims to make attraction over the premier league with a huge bulg of fans have the result of the match before the match with the accuracy of 94.81% .but from aside of the remaining 5.19% it may give a shock to the users and fans to by changing the result after all its conceder as the unpredictable turnament.

Rajesh Goel et al [02] the modal is developed by using the matches data of about 692 match from starting year of ipl to year 2016 (2007-2016) its near by the duration of 9 years in each year there are apx 60 + matches were conducted in various venues at different country's .

Daniel Mago Vistro, Faizan Rasheed, Leo Gertrude David et al [03] the aim of developing such a model is make attraction of users over fantasy game with an best accurate prediction in the game to get an high outcome from the game its similar to fun and earn

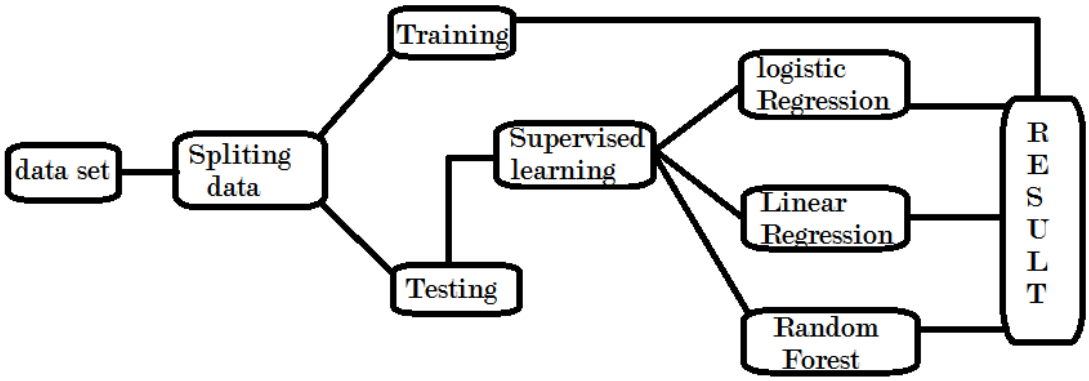
IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

Rabindra Lamsal & Ayesha Choudhary et al [04] the project and the model is based on 14 various fields mentioned in the data set like venue team a & b toss winner game winner and many other such 14 attributes which make it an better model to predict the result with high accuracy

III. PROPOSED SYSTEM

This system is developed to pridect the result of the game held in the Indian Premier League with the help of machine learning algorithm in jupyter notebook and colab (an google software or online platform to execute the code of ipynb format) have the data set provided by the kaggle.com .hear is an system model to predict the outcome as the winner of the match of an ipl with the accuracy of about 94.81 %.by the prediction base of data of previous 9 years (2007-2016).

SYSTEM ARCHITECTURE



IV. METHODOLOGY

DATA COLLECTION

We have collected the raw data of about 9 years of IPL history with total 692 matches with the venue match result and the other data from the kaggle.com who provide raw data of various thing and events in the form .csv folder .

	id	Season	city	date	team1	team2	toss_winne	toss_decisi	result	dl_applied	winner	win_by_rur	win_by_wi	player_of_	venue	umpire1	umpire2	umpire3
2	1	IPL-2017	Hyderabad	5/4/2017	Sunrisers H	Royal Chall	Royal Chall	field	normal	0	Sunrisers H	35	0	Yuvraj Singl	Rajiv Gandl	AY Dandeka	NJ Llong	
3	2	IPL-2017	Pune	6/4/2017	Mumbai Ind	Rising Pune	Rising Pune	field	normal	0	Rising Pune	0	7	SPD Smith	Maharashtr	A Nand Kisl	S Ravi	
4	3	IPL-2017	Rajkot	7/4/2017	Gujarat Lio	Kolkata Kni	Kolkata Kni	field	normal	0	Kolkata Kni	0	10	CA Lynn	Saurashtra	Nitin Meno	CK Nandan	
5	4	IPL-2017	Indore	8/4/2017	Rising Pune	Kings XI Pu	Kings XI Pu	field	normal	0	Kings XI Pu	0	6	GJ Maxwell	Holkar Cricl	AK Chaudh	C Shamshuddin	
6	5	IPL-2017	Bangalore	8/4/2017	Royal Chall	Delhi Daree	Royal Chall	bat	normal	0	Royal Chall	15	0	KM Jadhav	M Chinnaswamy Stadium			

DATA SEPRATION

We have to remove the unwanted data from the data set . while reading the data we found that the data stored in the data set is not full filled properly there are a few empty column in the dataset which make it complex to use in our model so for that we have to remove the empty column from data nad also the same or copy data from the dataset as we know that the data we have contain the team DD & DC are same but the model will consider them as a different unit in the data set so we have to merge them together

MODEL DEVELOPMENT

in this phase we are going to develop an model which will able to predict the winner of the match held in between two team of the ipl with the accuracy of 94.81% . we have used the basic ML algorithm to develop the model like leniar regression , seaborn , logic regression & other .

V. CONCLUSION

In this process, different Data Science methods were adopted to conduct the study, including data mining, visualization, preparation of database, feature engineering, applying the Analytic hierarchical process, creating prediction models, and

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
Impact Factor: 7.53
Volume 4, Issue 1, March 2024

Training classification techniques. The IPL dataset was gathered and pre-processed. The missing values were removed, and variables were encoded into the numerical format to make the dataset uniform. The important features were then derived from data using the domain knowledge to extract raw data features via data mining techniques, and the results were derived from the model. As the dataset that is available for IPL is limited and small, multiple levels of features were created to make sure that the derived model is not underfit. Almost every feature that can affect the result of a match was derived. A number of machine learning models were applied to the selected features to predict the IPL match results.

REFERENCES

[1]. J. Surowiecki and P. Silverman, "The wisdom of crowds", American Journal of Physics, Vol. 75, No. 2, pp. 190-192, 2005.

[2]. B. E. Sawe, "The Most Popular Sports in the World," *World Atlas*, 16-Sep-2016. [Online]. Available: https://www.worldatlas.com/articles/what-are-the-most-popular sports-in-the-world.html.

[3]. Kampakis, S. and Thomas, W. (2018). Using Machine Learning to Predict the Outcome of English County twenty over Cricket Matches. [online] Arxiv.org. Available at: https://arxiv.org/abs/1511.05837 [Accessed 1 Sep. 2018].

[4]. Bhandari, E. Colet, and J. Parker. Advanced Scout: Data mining and knowledge discovery in NBA data. Data Mining and Knowledge Discovery, 1(1):121–125, 1997.

[5]. Brooks, R. D., Faff, R. W., & Sokulsky, D. (2002). An ordered response model of test cricket performance. Applied Economics, 34 (18), 2353-2365.

[6]. N. Pathak and H. Wadhwa, "Applications of Modern Classification Techniques to Predict the Outcome of ODI Cricket", Procedia Computer Science, vol. 87, pp. 55-60, 2016.

[7]. P. Satao, A. Tripathi, J. Vankar, B. Vaje and V. Varekar, "Cricket Score Prediction System (CSPS) Using Clustering Algorithm", International Journal of Current Engineering and Scientific Research, vol. 03, no. 04, pp. 43-46, 2016.

[8]. P. Shah and M. Shah, "Predicting ODI Cricket Result", Journal of Tourism, Hospitality and Sports, vol. 05, pp. 19-20, 2015.

[9]. A. Kaluarachchi and S. V. Aparna, "CricAI: A classification based tool to predict the outcome in ODI cricket," 2010 Fifth International Conference on Information and Automation for Sustainability, Colombo, 2010, pp. 250-255.

[10]. M. Jhawar and V. Pudi, "Predicting the Outcome of ODI Cricket Matches: A Team Composition Based Approach", European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases, Riva del Garda, 2016