iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** X **Month of publication:** October 2023

DOI: https://doi.org/10.22214/ijraset.2023.56130

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue X Oct 2023- Available at www.ijraset.com

Cricket Win Prediction using Machine Learning

Sanika A. Jadhav[1], Shivendra M. Jadhav[2], Sahil M. Jadhav[3], Shubham M. Jadhav[4], Shreyash V. Jadhav[5], Sarvesh R. Jadhav[6], Supriya S. Telsang[7]
Department of Engineering, Sciences and Humanities (DESH) Vishwakarma Institute of Technology, Pune, 411037, Maharashtra, India

Abstract: ***Cricket, being one of the most popular sports worldwide, has attracted significant interest in developing accurate win prediction models. With the advent of machine learning techniques, researchers have leveraged the power of data-driven algorithms to predict cricket match outcomes. This research paper aims to improve cricket win prediction model by using XGBoost machine learning algorithm. Feature importance analysis is conducted to identify the most influential factors contributing to match outcomes. The dataset is divided into training and test sets, and the models are evaluated on both datasets to measure their generalization performance. The findings demonstrate the potential of machine learning techniques in accurately forecasting cricket match outcomes, enabling stakeholders to make informed decisions in the dynamic and unpredictable domain of cricket.***

Keywords: Machine Learning, Prediction, XGBoost, Cricket.

I. INTRODUCTION

In the realm of sports, cricket has always held a special place, with a fervent following and countless fan. Over time, a new dimension of cricket fandom has emerged through fantasy cricket leagues, transforming the way fans engage with the game. Fantasy cricket offers a unique opportunity for individuals to become virtual team owners and strategists, utilizing their knowledge of the sport to create winning combinations of players (Vistro, 2019).

With millions of enthusiasts participating in fantasy cricket leagues worldwide, the quest for gaining a competitive edge and maximizing team performance has become a paramount goal for players. To address this need, the application of machine learning algorithms in predicting fantasy cricket team wins has gained considerable attention. Fantasy cricket has emerged as a popular online gaming platform that allows users to create their virtual cricket teams and compete against each other based on the real-time performance of players in matches (Bhatia, 2020). Machine learning techniques, with their ability to analyse large volumes of data and identify patterns, have the potential to revolutionize the world of fantasy sports. By harnessing the power of historical player data, match statistics, pitch conditions, weather forecasts, and various other factors, machine learning models can provide valuable insights and predictions to aid fantasy cricket team selection (Basit, 2020).

Users are constantly on the lookout for tools and strategies that can help them assemble the most potent team, capable of outperforming opponents and securing victories. To cater to this demand, we present an innovative application that leverages the power of machine learning to predict fantasy cricket team wins.

II. METHODOLOGY/EXPERIMENTAL

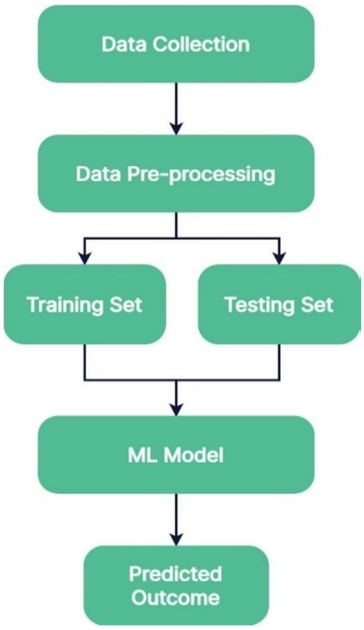
A. Problem Definition

The objective of this project is to accurately predict the winning team in fantasy game cricket by leveraging various factors. By analysing player statistics, match records, historical performance, pitch conditions, weather data, and other relevant variables, we aim to develop a machine learning model that can provide reliable predictions. The goal is to assist fantasy game cricket enthusiasts in making informed decisions when selecting their teams, increasing their chances of achieving higher scores and winning competitions.

B. Data Collection and Pre-processing

Then the next step is to collect the necessary data for training and evaluating the machine learning model. We gathered data from various sources, including cricket databases, match records, player statistics, and weather reports of Indian premier League which is Cricket League Played in India every year.

The Dataset include a wide range of variables such as inning, overs, ball number, batter, bowler, non-striker, extra type, batsman run, extra runs, total run, batting team, bowling team and other necessary parameter that are require to predict the output of model
Once the data was collected, we performed pre-processing to ensure its quality and suitability for analysis. We checked for and removed any duplicate records in the dataset to avoid biased results and inconsistencies.



C. Deciding Parameter and Hyperparameter

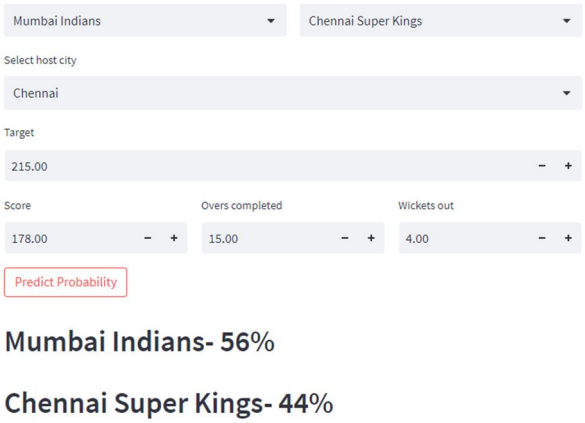
In the domain of fantasy game cricket, deciding on the appropriate parameters and hyperparameters is vital for constructing an effective prediction model. Parameters refer to the variables that the model learns during training, such as the weights and biases of a neural network. These parameters are adjusted iteratively to minimize the difference between predicted and actual outcomes, enhancing the model's predictive accuracy.
Hyperparameters, on the other hand, are set by the user before training the model and determine how the learning process takes place. In the context of predicting the winning team in fantasy cricket, several parameters and hyperparameters are key. Parameters such as inning, overs, ball number, batter, bowler, non-striker, extra type, batsman run, extra runs, total run, batting team, and bowling team provide essential information about the game dynamics, player performance, and team composition.
Hyperparameters, such as the learning rate, regularization strength, and model architecture, need to be carefully selected to optimize the model's performance. These choices influence how the model learns from the data, balances complexity and simplicity, and generalizes to new instances.
The selection of appropriate parameters and hyperparameters requires a deep understanding of the game, careful analysis of historical data, and iterative experimentation to fine-tune the model's predictive capabilities.

D. Training Machine learning Model

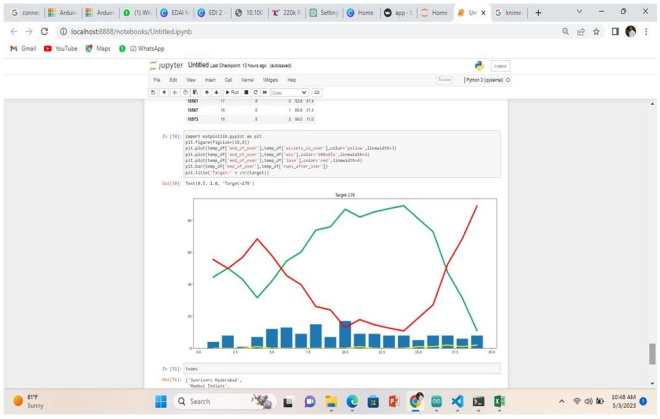
To train and evaluate the model, we divided the dataset into training and testing sets. The training set consists of data from 2008 to 2020, and the testing set comprises data from 2021 to 2022. This split allows us to assess the model's performance on unseen data, providing a realistic evaluation of its predictive capabilities.
Then model development approach described, which utilizes the Gradient Boosting algorithm, appears to be a suitable choice for building a win prediction model for fantasy game cricket teams. Gradient Boosting is known for its ability to create robust and accurate predictive models by combining multiple weak learners, such as decision trees.

III. RESULTS AND DISCUSSIONS

The XGBoost model achieved promising results in predicting cricket match outcomes. The accuracy and reliability of the model were evaluated using performance metrics such as accuracy, precision. The evaluation demonstrated the effectiveness of the model in capturing the dynamic nature of cricket matches and accurately predicting outcomes.



The results showcase the potential of XGBoost in the field of sports analytics and its ability to contribute to decision-making processes in cricket. Future research can explore the integration of additional features and the development of ensemble models to further improve prediction accuracy and robustness. The successful implementation of the XGBoost algorithm in predicting cricket match outcomes. The developed Streamlit web application provides cricket enthusiasts and professionals with a user-friendly interface for obtaining real-time predictions during ongoing matches.



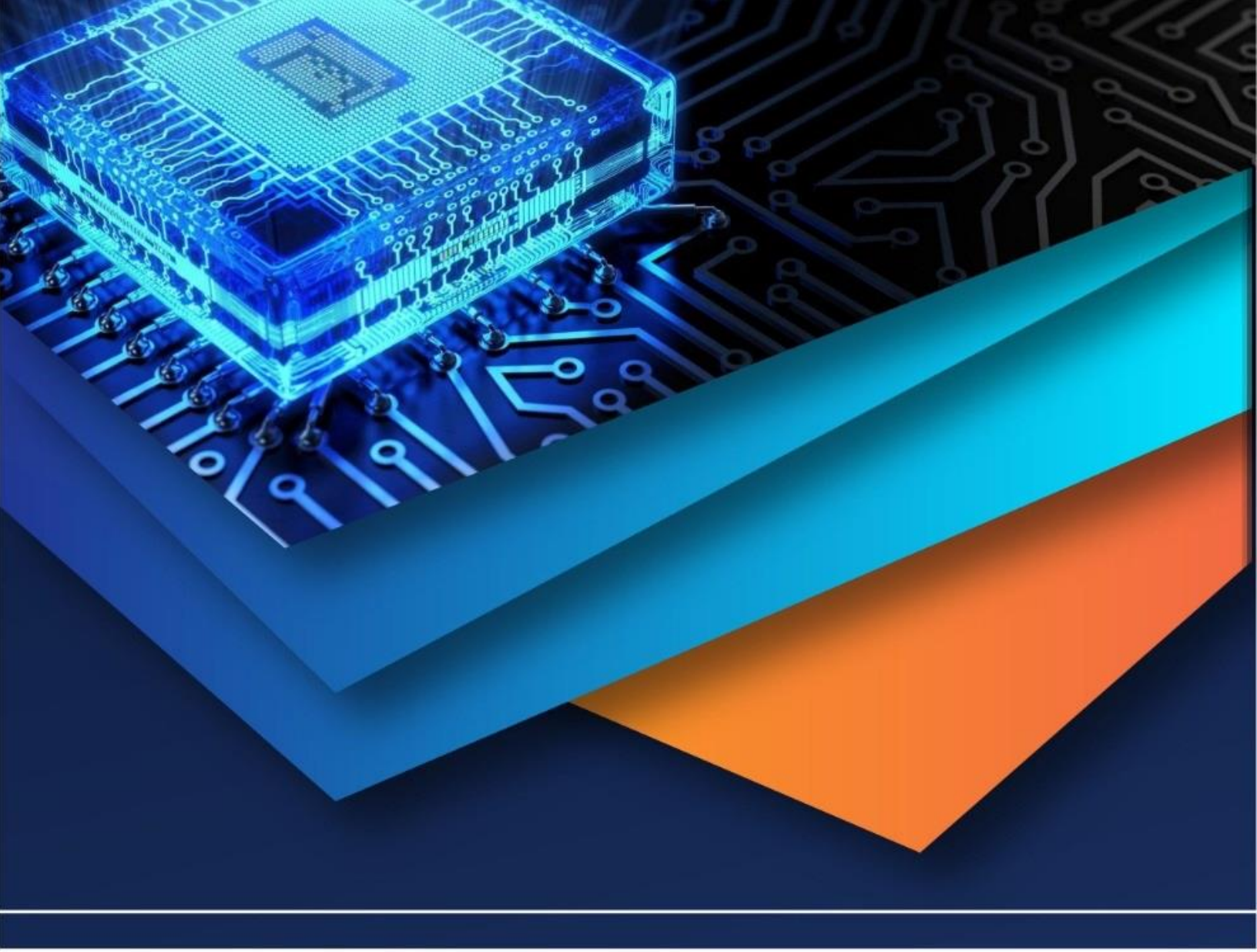
Logistics Regression (existing algorithm)	XGboost Algorithm (current algorithm)
1.The accuracy of the testing data on the logistic regression model is 80%. 2. Simple and interpretable model, with coefficients indicating the relationship between input variables and the outcome.	1.The accuracy of the testing data on the XGboost algorithm is 99%. 2. the XGBoost method has better results based on four evaluation indicators namely accuracy, sensitivity, specificity, and precision.

IV. ACKNOWLEDGEMENT

We would also like to thank the staff and students of the Department of Sciences and Humanities at Vishwakarma Institute of Technology for their valuable feedback and support. We would also like to extend our appreciation to the reviewers who provided constructive comments that helped us improve the quality of this paper.

REFERENCES

[1] Vistro, Daniel Mago, Faizan Rasheed, and Leo Gertrude David. "The cricket winner prediction with application of machine learning and data analytics." International Journal of Scientific & Technology Research 8.09 (2019).
[2] Pramanik, Md Aktaruzzaman, et al. "Performance Analysis of Classification Algorithms for Outcome Prediction of T20 Cricket Tournament Matches." 2022 International Conference on Computer Communication and Informatics (ICCCI). IEEE, 2022.
[3] WEERADDANA, NIMMI, and SAMINDA PREMARATNE. "Unique approach for cricket match outcome prediction using xgboost algorithms." Journal of Theoretical and Applied Information Technology 99, no. 9 (2021): 2162-2173.
[4] Srivastava, P. R., Eachempati, P., Kumar, A., Jha, A. K., & Dhamotharan, L. (2022). Best strategy to win a match: an analytical approach using hybrid machine learning-clustering-association rule framework. Annals of Operations Research, 1-43.
[5] Kapadia, Kumash, Hussein Abdel-Jaber, Fadi Thabtah, and Wael Hadi. "Sport analytics for cricket game results using machine learning: An experimental study." Applied Computing and Informatics ahead-of-print (2020).
[6] Ishi, Manoj, D. J. Patil, D. N. Patil, and D. V. Patil. "Winner prediction in one day international cricket matches using machine learning framework: An ensemble approach." Indian Journal of Computer Science and Engineering 13 (2022): 628-641.
[7] MOUNIKA, POTHNAK, MADDALA NIKHILA DEVI, JAVERIA NISHAT, and GADDAM LIKHITA. "IPL WINNER PREDICTION USING MACHINE LEARNING." Journal of Engineering Sciences 13, no. 7 (2022).
[8] Bhatia, Vandana. "A review of Machine Learning based Recommendation approaches for cricket." In 2020 Sixth International Conference on Parallel, Distributed and Grid Computing (PDGC), pp. 421-427. IEEE, 2020.
[9] Shakil, F. A., Abdullah, A. H., Momen, S., & Mohammed, N. (2020). Predicting the Result of a Cricket Match by Applying Data Mining Techniques. In Software Engineering Perspectives in Intelligent Systems: Proceedings of 4th Computational Methods in Systems and Software 2020, Vol. 2 4 (pp. 758-770). Springer International Publishing.
[10] Basit, Abdul, Muhammad Bux Alvi, Fawwad Hassan Jaskani, Majdah Alvi, Kashif H. Memon, and Rehan Ali Shah. "ICC T20 Cricket World Cup 2020 winner prediction using machine learning techniques." In 2020 IEEE 23rd International Multitopic Conference (INMIC), pp. 1-6. IEEE, 2020.
[11] Singh, Tejinder, Vishal Singla, and Parteek Bhatia. "Score and winning prediction in cricket through data mining." 2015 international conference on soft computing techniques and implementations (ICSCTI). IEEE, 2015.
[12] Pallavi Tekade, Kunal Markad, Aniket Amage and Bhagwat Natekar (2020). “CRICKET MATCH OUTCOME PREDICTION USING MACHINE LEARNING”and Bhagwat Natekar (2020). “CRICKETMATCH OUTCOME PREDICTION USINGMACHINE LEARNING”
[13] Dhonge, N., Dhole, S., Wavre, N., Pardakhe,M., & Nagarale, A (2021). “IPL Cricket Score and Winning Prediction Using Machine Learning Techniques”
[14] “Cricket Analysis and Prediction of projected Score and Winner using Machine Learning” Akhil Nimmagadda, Nidamanuri Venkata Kalyan, Manigandla Venkatesh, Nuthi Naga Sai Teja, Chavali Gopi Raju (2018). “Cricket score and winning prediction using data mining”
[15] A. Bandulasiri, “Predicting the winner in one day international cricket,” Journal of Mathematical Sciences & Mathematics Education, vol. 3, no. 1, pp. 6–17, 2008
[16] Nimmagadda, Akhil, et al. "Cricket score and winning prediction using data mining." International Journal for Advance Research and Development 3.3 (2018): 299-302
[17] Dhonge, N., Dhole, S., Wavre, N., Pardakhe, M., & Nagarale, A. IPL CRICKET SCORE AND WINNING PREDICTION USING MACHINE LEARNING TECHNIQUES. https://irjmets.com/uploadedfiles/paper/volume3/issue_5_may_2021/10362/1628083416.pdf
[18] Ahmed, W. & Nazir, K., 2015. A Multivariate Data Mining Approach to Predict Match Outcome in One-Day International Cricket. 10.13140/RG.2.2.30683.46880.
[19] Hossin, M. & Sulaiman, M., 2015. A REVIEW ON EVALUATION METRICS FOR DATA CLASSIFICATION EVALUATIONS. International Journal of Data Mining & Knowledge Management Process (IJDKP) , 5(2)
[20] Bunker, Rory & Thabtah, Fadi. (2017) “A Machine Learning Framework for Sport Result Prediction.
[21] Applied Computing and Informatics”, 15. 10.1016/j.aci.2017.09.005.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY