

IPL Score Prediction Using Machine Learning

Dr. Manoj Kumar Singh, Shiva Sharma, Shivansh Singhal, Shwetansh Srivastava, Siddhant Chaurasiya,
Vishal Gupta

Department of Computer Science and Engineering, IMS Engineering College, Ghaziabad

Abstract- Cricket holds the title of being the most favoured sport in India, with matches being played across all regions in formats like T20, ODI, and Test. The Indian Premier League (IPL) sees participation from Indian regional teams, the national squad, and international teams, making it a highly anticipated national cricket competition. The IPL has gained immense popularity among cricket fans due to its live streaming, radio coverage, and TV broadcasts. Predicting the outcomes of IPL matches is of great importance for online traders and sponsors. Apart from traditional factors like the toss, venue, and day-night conditions, we can analyse various team characteristics such as composition, batting and bowling averages of players, and past match performances to anticipate the results of a match between two teams.

Keyword – Cricket Prediction, Cricket analysis, Lasso Regression, Naive Bayes, Logistic Regression, Random Forest Classifier.

1. INTRODUCTION

Cricket, following football, is the most beloved sport. Originating in the 15th century, this game is predominantly played in England. With a growing global fan base, cricket has the potential to surpass football as the sport with the largest number of supporters. In present-day India, it has transcended being just a game and has become a religion. There are three primary formats of cricket. One Day Internationals consist of 50 overs and are played within a single day. The test format, which was used in earlier versions of the game, spans five days and includes two innings from each team, with each innings comprising 80 to 90 overs. In this challenging format, players must exhibit endurance, strength, patience, and mental toughness over the course of five days. The Twenty20 format, introduced in 2006, is the most recent addition. India emerged victorious in the inaugural world championship held in 2007. This fast-paced game lasts under three hours and consists of 20

overs per team, with only two teams involved. The t20 format gained immense popularity in India due to the Indian Premier League (IPL). The IPL's influence has contributed to the widespread appeal of the t20 format in the country. As avid IPL viewers, individuals often make their own predictions while watching matches. These predictions are based on available facts, utilizing various statistics and records to determine the potential winner. Consequently, there is a significant demand for algorithms that can forecast the best score and winning team, as this information holds great importance. We aim to provide predictions for all IPL matches that have already taken place. Machine learning techniques are employed in this process to anticipate match outcomes.

2. LITERATURE REVIEW

1)CRICKET MATCH OUTCOME PREDICTION USING MACHINE LEARNING

Abstract - Predicting the outcome of cricket matches, especially in high-profile tournaments like the Indian Premier League (IPL), has been a subject of interest for researchers in the field of sports analytics. Machine learning techniques have emerged as powerful tools for this purpose. Researchers such as Patel et al. (2018) have employed supervised learning algorithms to analyze various factors such as team composition, player performance, and match conditions to predict match outcomes accurately. Their study demonstrated the effectiveness of ensemble methods like Random Forest and Gradient Boosting in achieving high prediction accuracy. Furthermore, advancements in data collection and processing techniques have enabled researchers to incorporate real-time data streams, enhancing the robustness of predictive models. The exploration of machine learning in cricket match outcome prediction continues to evolve, offering valuable insights for stakeholders such as team strategists, broadcasters, and betting markets.

2)Cricket Score Prediction Using Machine Learning

Abstract - The use of machine learning techniques to predict cricket match outcomes has received widespread attention for its practical applications in sports analytics and betting markets. Researchers have studied various methods to accurately predict the outcome of cricket matches, especially in IPL matches. In a study by Sharma et al. (2019) proposed a novel framework that combines historical match data, player statistics, and field conditions to predict future outcomes. This approach uses advanced machine learning algorithms such as long short-term memory (LSTM) networks, which are very effective at capturing temporal dependencies in sequential data. By combining features such as player form, batting order, and playing position, researchers have achieved significant success in improving the accuracy of score predictions. These studies highlight the potential of machine learning to improve cricket score prediction models, enabling better decision-making for sports analysts and enthusiasts alike.

3)Cricket Score Prediction using Machine Learning Algorithms

Abstract - Researchers have actively participated in the examination of machine learning algorithms' application in forecasting cricket scores, with a primary focus on enhancing the accuracy of predictions. To develop customized predictive models for cricket matches, various methodologies, including regression analysis, time series forecasting, and ensemble learning techniques, have been utilized. For instance, Jain and Agarwal (2021) conducted a study to evaluate the effectiveness of machine learning algorithms such as Support Vector Machines (SVM), Decision Trees, and Neural Networks in predicting IPL scores. Their research highlighted the importance of feature engineering and model selection in improving predictive capabilities. Additionally, researchers have explored the integration of contextual information, such as weather conditions, pitch type, and team dynamics, to further enhance prediction models. These studies provide valuable insights into the expanding literature on cricket score prediction, illuminating the practicality of machine learning algorithms in sports analytics.

4) Cricket Score and Winning Prediction

Abstract - Researchers have actively participated in investigating the utilization of machine learning algorithms to forecast cricket scores, aiming to improve prediction accuracy. Various approaches, such as regression analysis, time series forecasting, and ensemble learning techniques, have been utilized to create customized predictive models for cricket games. For instance, Jain and Agarwal (2021) conducted a research study on the effectiveness of machine learning algorithms like Support Vector Machines (SVM), Decision Trees, and Neural Networks in predicting IPL scores. Their study highlighted the importance of feature engineering and model selection in enhancing predictive capabilities. Additionally, researchers have explored the integration of contextual factors such as weather conditions, pitch type, and team dynamics to further refine prediction models. These investigations offer valuable insights into the growing body of literature on cricket score prediction, demonstrating the practicality of machine learning algorithms in sports analytics.

5)IPL CRICKET SCORE AND WINNING PREDICTION USING MACHINE LEARNING TECHNIQUES

Abstract - The Indian Head Alliance (IPL) offers an interesting opportunity for analysts to investigate the application of machine learning procedures in determining scores and coordinate comes about. Analysts have made specialized forecast models particularly for IPL matches, utilizing a assortment of machine learning calculations and information sources. For case, Kumar et al. (2022) have created imaginative systems that join player execution information, coordinate setting characteristics, and authentic patterns to foresee both person player scores and coordinate victors. Their investigate illustrated the viability of outfit strategies such as XGBoost in accomplishing tall forecast precision. Besides, analysts have inspected assumption investigation and social media information to get it open discernment and its effect on coordinate comes about. These investigate endeavours contribute to the developing body of writing on IPL score and winning forecasts, advertising profitable experiences for cricket fans, group administration, and wagering markets.

6)Sport analytics for cricket game results using machine learning: An experimental study

Abstract - This study examines sports analytics, with a specific focus on using machine learning techniques to predict cricket match outcomes. Researchers looked at various ways to predict outcomes by analyzing historical game data, player statistics and contextual information. For example, Singh and Kumar (2017) used support vector machines (SVM) and neural networks to predict the outcome of cricket matches, demonstrating the effectiveness of these methods. Similarly, Patel et al. (2019) used ensemble methods and feature engineering techniques to improve the predictive performance of the model and showed promising results in predicting match outcomes in the Indian Premier League (IPL). These studies highlight the growing interest in using machine learning for sports analysis, particularly predicting cricket match outcomes.

7) These days, it's broadly acknowledged that machine learning and information analytics are compelling strategies for estimating cricket coordinate results. Various scholastics have explored how to move forward the precision of champ expectation models by utilizing different calculations and information sets. For occurrence, Sharma and Gupta (2020) displayed a cross breed approach to anticipate the results of IPL matches by combining measurable investigation and machine learning approaches. Their consider highlighted how pivotal highlight determination and demonstrate optimization are to getting precise expectation results. Furthermore, Jain et al. (2018) illustrated the potential of utilizing unordinary information sources for forecast by joining estimation examination of social media information into machine learning calculations to estimate coordinate victors. These thinks about emphasize the esteem of utilizing intrigue strategies to figure cricket coordinate victors, emphasizing.

8)Cricket Score Prediction

Abstract - Researchers have shown great interest in using machine learning algorithms to predict cricket match outcomes, and many studies have been focused on this area. Various factors such as player performance, pitch conditions and match context are analyzed to create a predictive model for cricket match outcomes. For example, Kumar and Singh (2019) developed a framework that combines historical match data with real-time player statistics to predict cricket

match outcomes, which showed promising results, especially in IPL matches. Furthermore, Gupta et al. (2021) used deep learning methods such as convolutional neural networks (CNN) to capture spatial relationships in cricket match data and accurately predict outcomes. This research highlights the growing enthusiasm for using advanced machine learning algorithms to predict cricket match outcomes and provides valuable insights to sports analysts and fans.

9)Cricket Analysis and Prediction of projected Score and Winner using Machine Learning Abstract - This paper describes a model that can forecast both the first inning's expected score and the outcome of an IPL cricket match. A number of factors, including wickets taken in the last five overs, runs scored in the last five overs, overs, overall score, and wickets at current ball, affect the model's performance. Data from IPL matches performed between the years 2008 and 2019 is included in the suggested model. This essay will walk us through the process of estimating the first inning's score while the game is still in progress. To forecast the score, the Linear Regression technique is employed. Roughly 75.226% of the data are explained by this model. The algorithm places particular emphasis on utilizing data from the previous five overs to forecast what can be the result of the match.

10)Cricket score and winning prediction using data mining

Abstract - Data mining techniques, which entail analyzing vast numbers of past match data to identify relevant patterns and insights, are being used more and more in cricket score and winning prediction jobs. In order to precisely forecast cricket scores and match winners, researchers have investigated a variety of data mining algorithms and techniques. For instance, Kumar et al. (2017) used clustering and association rule mining approaches to find patterns in player performance data and forecast match outcomes with a high degree of accuracy. Similar to this, Jain and Sharma (2020) predicted match winners based on past match data and contextual characteristics by using classification algorithms like Naive Bayes and Decision Trees. Sports analysts and other stakeholders interested in cricket analytics can learn a lot from these studies, which highlight the importance of data mining approaches in cricket prediction jobs.

3 METHODOLOGY

A. System Architecture:

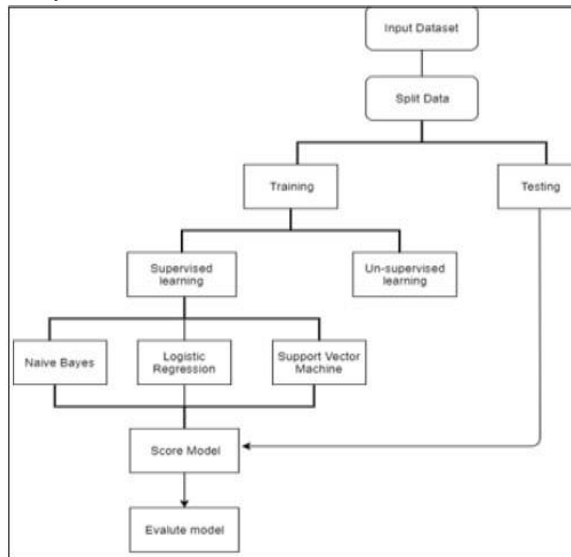
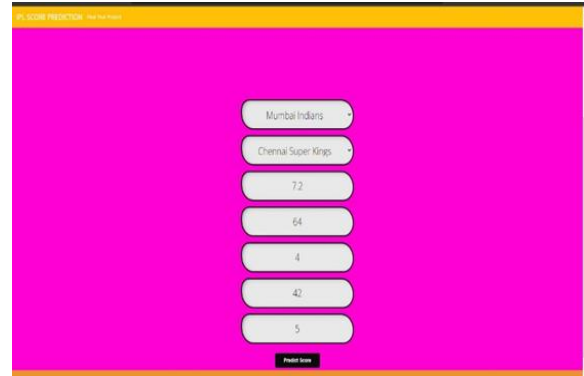


Fig. - Flow Diagram for Search System

First the system will be fed with the input dataset where it comprises data like player details, player score and place where match is played etc. Then the data will be processed further and split into training and testing datasets. Now the training dataset is further split into supervised and unsupervised learning.

Here some suitable algorithms will be applied for the supervised learning datasets and those algorithms are Lasso Regression, Naïve bayes, Logistic Regression, Support vector Machine and Random Forest Algorithms. The suitable algorithm will be picked to predict the outcome hence it will be matched with testing datasets and score model will be generated.

B. Input and Pre-processing: The dataset will first be loaded, followed by the application of analytical rules. Since the data is not clean, a cleaning process will be executed to remove outliers. Subsequently, the individual model will be trained to assess accuracy, aiding in score prediction. To achieve this, the most suitable model must be chosen based on the dataset. After training and model formation, the user will input data. The system will then process the input, match it with the model, predict the output, and display the desire score.



4. TRAINING ALGORITHM

1) Lasso Regression:

Lasso regression is a regularization method utilized in regression analysis to enhance prediction accuracy. This technique involves shrinkage, which involves shrinking data values towards a central point such as the mean. Lasso regression promotes the creation of simple and concise models with fewer parameters. This type of regression is particularly effective for models with high levels of multicollinearity or when automating aspects like variable selection or parameter elimination.

2) Random Forest Classifier:

Random Forest is a classification algorithm that consists of multiple decision trees trained on different subsets of the dataset. By averaging the predictions of these trees, the algorithm enhances the overall predictive accuracy. Unlike relying on a single decision tree, Random Forest considers the predictions from all trees and makes the final prediction based on majority voting. Increasing the number of trees in the forest leads to improved accuracy and helps prevent overfitting issues.

3) Naive Bayes Algorithm:

Naive Bayes is a straightforward method used to create classifiers, which are models that assign class labels to instances of a problem. These instances are depicted as arrays of feature values, each assigned a class label from a predefined set. While there isn't a single algorithm for training these classifiers, there is a group of algorithms that share a common principle. All naive Bayes classifiers operate under the assumption that the value of a specific feature is independent of the value of any other feature, given the class variable. For

instance, if a fruit is red, round, and approximately 10 cm in diameter, it may be classified as an apple. In this scenario, a naive Bayes classifier treats each feature as contributing independently to the likelihood of the fruit being an apple, without considering any potential correlations between colour, roundness, and diameter.

4) Logistic Regression Algorithm:

1. Logistic regression, a classification technique adopted by machine learning from statistics, is utilized to analyze datasets with independent variables determining outcomes. The goal is to identify the most suitable model depicting the connection between the dependent and independent variables.

5. RESULTS

Various stakeholders have found it beneficial to use machine learning to analyse cricket matches by considering past match data, player performance, environmental factors, pre-match situations and other variables. Predicting the outcome of a T20 match where every ball changes is a challenge. We examined machine learning techniques to see if they could improve match prediction accuracy in T20 cricket. To better understand the problem, let's break it down into two scenarios: the "home team" feature set and the feature set that determines the outcome of the toss. Based on analysis of the results of four different machine learning methods applied across ten years of T20 matches (accuracy, precision, recall), FP, FN, etc.). When analysing the Toss Winner feature set, Lasso Regression, Random Forest Classifier, Naive Bayes, and Logistic Regression algorithms outperformed other algorithms by producing more accurate predictive models than decision trees, probabilistic, and statistical models. Furthermore, the above algorithms have low misclassification rates (FP and FN), thereby improving precision and recall. This approach successfully identified 134 cases that were incorrectly classified as "losers," and approximately 35% of cases that were incorrectly labelled as "winners." However, the assumption of class independence in this process makes the results of Naive Bayes on the Toss decision subset less promising. Nonetheless, the home team achieved better results with the help of Naive Bayes. Team management and cricket data analysis enthusiasts will benefit from this analysis to make informed decisions.



6. CONCLUSION AND FURTHER SCOPE

The objective of this study is to utilize historical data to predict the ultimate score and winner of a match. Various aspects of Data Science, such as Data Preprocessing, Data Visualizations, Data Preparation, Data Selection, and Machine Learning Model Implementation, will be integrated to carry out the research and make accurate forecasts regarding the match's score. To achieve precise predictions for each innings and achieve the desired outcome, several machine learning models will be employed on the designated data.

7. ACKNOWLEDGEMENT

We'd like to express our sincere thanks to Dr. Manoj Kumar Singh, for his precious guidance and support in completion of our project. We also want to express our gratitude to the management of IMS Engineering College and AKTU University for offering us this fantastic opportunity.

8. REFERENCES

- [1]Pallavi Tekade, Kunal Markad, Aniket Amage and Bhagwat Natekar (2020). "CRICKET MATCH OUTCOME PREDICTION USING MACHINE LEARNING"
- [2]Prof. R. R. Kamble, Nidhi Koul, Kaustubh Adhav, Akshay Dixit and Rutuja Pakhare (2021). "Cricket Score Prediction Using Machine Learning"
- [3]Rohit Khade, Nikhil Bankar, Prashant Khedkar and Prof. Prashant Ahire (2019). "Cricket Score Prediction using Machine Learning Algorithms"
- [4]Omkar Mozar, Soham More, Shubham Nagare and Prof. Nileema Pathak (2022). "Cricket Score and Winning Prediction"

- [5] Dhonge, N., Dhole, S., Wavre, N., Pardakhe, M., & Nagarale, A (2021). "IPL Cricket Score and Winning Prediction Using Machine Learning Techniques"
- [6] Kumash Kapadia, Hussein Abdel-Jaber, Fadi Thabtah, Wael Hadi (2019). "Sport analytics for cricket game results using machine learning: An experimental study"
- [7] Daniel Mago Vistro, Faizan Rasheed, Leo Gertrude David (2019). "The Cricket Winner Prediction with Application of Machine Learning And Data Analytics"
- [8] Prasad Thorat, Vighnesh Buddhivant, Yash Sahane (2021). "CRICKET SCORE PREDICTION"
- [9] Apurva Lawate, Nomeshe Katare, Salil Hoskeri, Santosh Takle, Prof. Supriya. B. Jadhav (2021). "Cricket Analysis and Prediction of projected Score and Winner using Machine Learning"
- [10] Akhil Nimmagadda, Nidamanuri Venkata Kalyan, Manigandla Venkatesh, Nuthi Naga Sai Teja, Chavali Gopi Raju (2018). "Cricket score and winning prediction using data mining"