iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 14 **Issue:** II **Month of publication:** February 2026

DOI: https://doi.org/10.22214/ijraset.2026.77395

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 14 Issue II Feb 2026- Available at www.ijraset.com

IPL Analysis Using Machine Learning

Chandra Shekar S[1], Abhirami P P[2]
Assistant professor HKBK Degree College, Nagavara, Bangalore-560045

Abstract: ***The Indian Premier League (IPL), as one of the most popular Twenty-20 cricket leagues worldwide, generates vast amounts of data related to matches, players, teams, and playing conditions, making it an ideal domain for machine learning–based analysis. This paper presents a comprehensive study on the application of machine learning techniques for analyzing IPL data with the objectives of predicting match outcomes, evaluating player performance, and supporting strategic decision-making. Using historical IPL datasets comprising match details, player statistics, team compositions, pitch conditions, and weather factors, the data is preprocessed and transformed through feature engineering to enhance predictive capability. Various supervised machine learning algorithms, including regression, classification, ensemble, and deep learning models, are employed to forecast match results and performance metrics. The study builds upon existing research that demonstrates the effectiveness of models such as Random Forest, Support Vector Machines, and deep learning architectures in sports analytics. The results highlight the potential of machine learning to uncover meaningful patterns, improve prediction accuracy, and provide actionable insights for stakeholders such as team management, analysts, fantasy sports users, and broadcasters. Overall, the proposed approach emphasizes the growing role of data-driven methods in modern cricket analytics and their contribution to enhancing competitiveness, audience engagement, and informed decision-making in the IPL ecosystem.***

I. INTRODUCTION

A. Overview of Machine Learning

Machine Learning is a branch of Artificial Intelligence that aims at solving real-life engineering problems. This technique requires no programming, whereas it depends on only data learning where the machine learns from pre-existing data and predicts the result accordingly. Machine Learning methods have benefit of using decision trees, heuristic learning, knowledge acquisition, and mathematical models. It thus provides controllability, observability, stability and effectiveness. Cricket is being played in many countries around the world. There are a lot of domestic and international cricket tournaments being held in many countries. The cricket game has various formssuch as Test Matches, Twenty20 Internationals, Internationals one day, etc. IPL is also one of them,and has great popularity among them. It's a twenty-20 cricket game league played to inspire youngand talented players in India. The league was conducted annually in March, April or May and has a huge fan base among India. There are eight teams which represent eight cities which are chosen from an auction. These teams compete against each other for the trophy. The whole match dependson the luck for the team, player's performance and lot more parameters that will be taken in to theconsideration. The match that is played before the day is also will make a change in the prediction. The stakeholders are much more benefited due to the huge popularity and the huge presence of people at the venue. The accuracy of a data depends on the size of the data we take for analysing and the records that are taken for predicting the outcome. Cricket is a game played between two teams comprising of 11 players in each team. The result is either a win, loss or a tie. However, sometimes due to bad weather conditions the game is also washed out as Cricket is a game which cannot be played in rain. Moreover, this game is also extremely unpredictable because at every stage of the game the momentum shifts to one of the teams between the two. A lot of times the result gets decided on the last ball of the match where thegame gets really close. Considering all these unpredictable scenarios of this unpredictable game, there is a huge interest among the spectators to do some prediction either at the start of the game orduring the game. Many spectators also play betting games to win money.

B. Problem Statement

The Indian Premier League (IPL) has emerged as one of the most popular and lucrative cricket leagues worldwide, attracting top players and massive audiences. Analyzing IPL matches using machine learning techniques presents an exciting opportunity to delve into the intricate dynamics of the game, player performances, team strategies, and match outcomes. The problem statement revolves around developing predictive models that can accurately forecast match results, player performances, and other key metrics based on various factors such as historical match data, player statistics, pitch conditions, weather forecasts, and team compositions. These models can aid team management, fantasy cricket enthusiasts, broadcasters, and betting platforms in making informed decisions and predictions.

©IJRASET: All Rights are Reserved

To address this problem, the analysis will involve preprocessing and cleaning large volumes of IPL data to extract meaningful features. Feature engineering will play a crucial role in selecting and transforming relevant attributes that contribute to the predictive power of the models. Machine learning algorithms such as regression, classification, clustering, and ensemble methods will be employed to train models on historical IPL data. Techniques like cross-validation, hyperparameter tuning, and model evaluation metrics will be utilized to ensure the robustness and generalization of the models. The ultimate goal is to develop accurate and interpretable models that can provide valuable insights into IPL matches, enabling stakeholders to make data-driven decisions in the dynamic and competitive landscape of professional cricket.

C. Objective of the Project

The objective of employing machine learning techniques in the analysis of the Indian Premier League (IPL) is to leverage data-driven insights for improving decision-making, performance evaluation, and competitive outcomes across the league. By utilizing historical match data, player statistics, pitch and weather conditions, and team compositions, machine learning models can accurately predict match results, enabling cricket enthusiasts, fantasy league participants, and betting platforms to make informed decisions. Additionally, machine learning facilitates in-depth player performance analysis by predicting individual contributions based on past and current form, supporting team management in player selection, strategy formulation, and talent identification. These techniques also help optimize team strategies by uncovering patterns in batting orders, bowling rotations, and fielding setups that contribute to successful outcomes. Furthermore, machine learning enhances audience engagement and broadcasting experiences through real-time analytics, predictive insights, and interactive visualizations during live matches. Overall, the application of machine learning in IPL analysis supports strategic optimization, enriches fan experience, and fosters continuous player development, thereby driving innovation and competitiveness in professional cricket.

II. LITERATURE REVIEW (Short version)

Several studies have explored the application of machine learning techniques for predicting match outcomes, scores, and player performance in the Indian Premier League (IPL). Kiran Gawande et al. [1] utilized classification algorithms such as Decision Tree, Naive Bayes, and KNN on historical IPL data to predict match results, highlighting the role of feature engineering and data preprocessing. Nikhil Dhonge et al. [2] focused on both score and winner prediction using regression and classification models, concluding that Random Forest and Linear Regression yield superior performance. Sudhamathy et al. [3] compared multiple classifiers and identified Random Forest as the most accurate model for predicting match winners. Deep learning approaches were introduced by Priyanka Kumbhar et al. [4], who employed RNN and LSTM models for IPL score prediction, demonstrating their effectiveness in handling sequential cricket data. Archana Kumar [5] and G. V. Gayathri [6] emphasized regression-based score prediction and classification-based winner prediction, reporting strong performance from Random Forest and SVM models. T. A. Severini [7] explored multiple supervised learning algorithms and showed improved accuracy over earlier models by selecting optimal features. Kilari Jaswanth et al. [8] highlighted the importance of data exploration and visualization while applying machine learning techniques to analyze player and team performance, with Random Forest achieving the highest accuracy. Overall, the literature indicates that machine learning, particularly ensemble and deep learning methods, plays a significant role in enhancing prediction accuracy and strategic insights in IPL analytics

III. METHODOLOGY

Data collection is a critical step in IPL analysis using machine learning, as the accuracy and reliability of predictions largely depend on the quality of the data gathered. The dataset comprises comprehensive information related to IPL matches, including match details such as dates, venues, participating teams, toss results, innings scores, and final outcomes. In addition, detailed player statistics are collected, covering batting, bowling, and fielding performances to analyze individual contributions and predict future performance. Team composition data, including playing elevens, batting and bowling orders, and lineup changes, is also considered to understand team strategies. Furthermore, pitch conditions and weather parameters such as temperature, humidity, rainfall, and wind speed are incorporated, as these factors significantly influence match dynamics. Additional data such as player demographics, historical season records, and team rankings may also be included based on analysis requirements. The data is sourced from reliable platforms such as official IPL records, cricket statistics databases, publicly available datasets, and weather forecasting services. To ensure effective model training, the collected data undergoes preprocessing steps including cleaning, normalization, and feature engineering, while maintaining data accuracy, consistency, and ethical compliance.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

IV. MACHINE LEARNING ALGORITHMS

In IPL analysis, a variety of machine learning algorithms are employed to predict match outcomes, analyze player performance, and support strategic decision-making. Linear regression is used for predicting continuous variables such as total runs scored by a team or individual player, helping to understand the relationship between performance metrics and match conditions. Logistic regression is applied to binary classification problems, such as predicting match results or specific player achievements, by estimating outcome probabilities based on input features. Decision tree algorithms are effective in handling both categorical and numerical data, making them suitable for identifying key factors influencing match outcomes and player performance. Random Forest, an ensemble of decision trees, enhances prediction accuracy and reduces overfitting, and is widely used for match result prediction and feature importance analysis. Support Vector Machines (SVM) are employed for classification tasks by determining optimal decision boundaries between classes. K-Nearest Neighbor (KNN) algorithms predict outcomes based on similarity measures, classifying instances according to their nearest neighbors in the dataset. Additionally, deep learning techniques such as artificial neural networks and convolutional neural networks are utilized for capturing complex patterns and nonlinear relationships in large IPL datasets, enabling more accurate predictions of match outcomes, player performances, and team strategies.

V. CONCLUSION

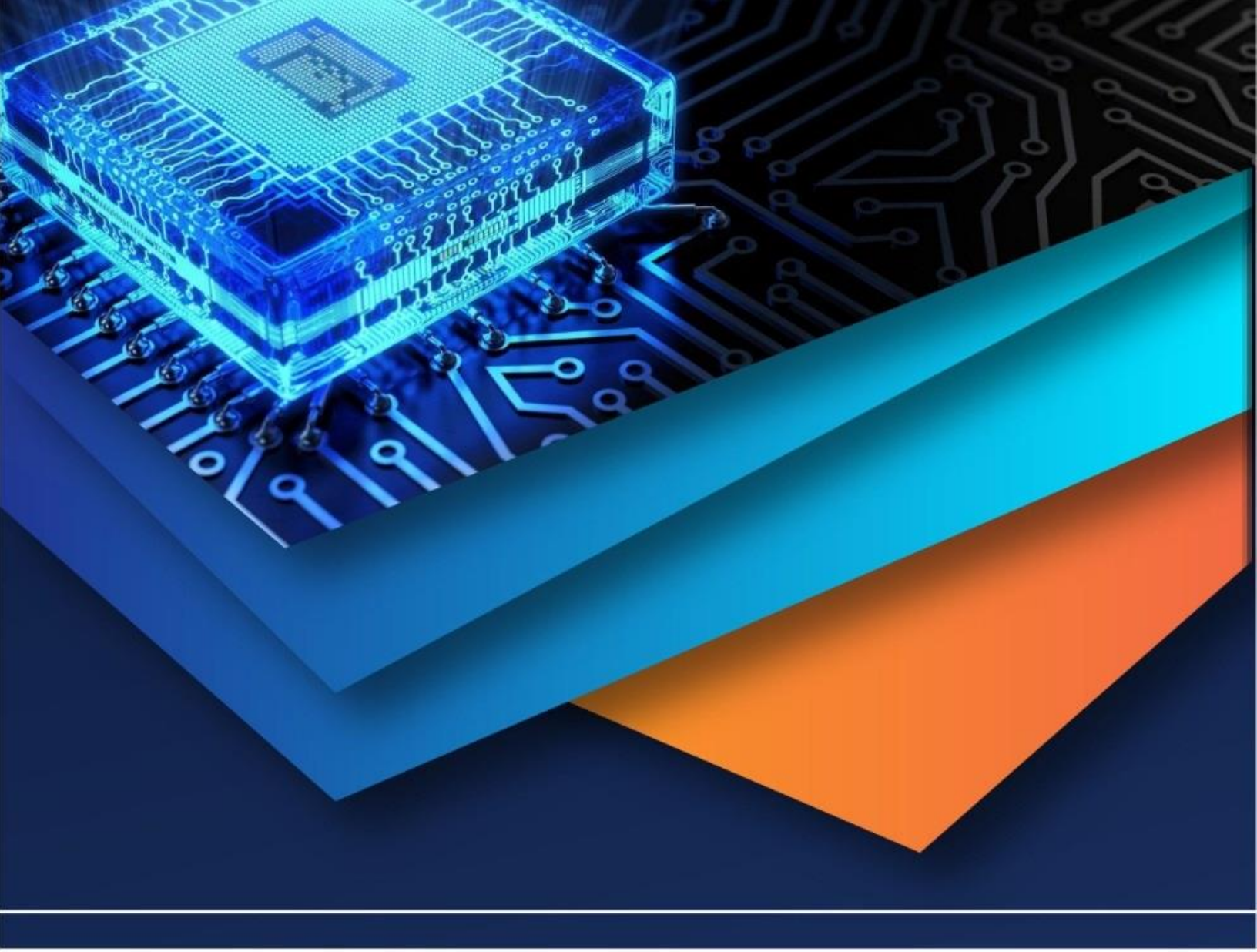
Utilizing AI for IPL investigation offers a significant comprehension of player elements, group systems, and match results. Through fastidious information assortment, preprocessing, and include designing, combined with the use of different AI models, this investigation discloses noteworthy bits of knowledge for partners. From anticipating player exhibitions to determining match results, the models give important direction to group the board, mentors, and players. The assessment of model execution highlights the unwavering quality and precision of the expectations, engaging IPL partners with informed dynamic capacities. Also, the translation of model results reveals insight into basic patterns, qualities, shortcomings, and regions for development inside groups. By making an interpretation of information driven experiences into noteworthy suggestions, this examination works with vital changes and upgrades the upper hand of IPL groups in ongoing seasons. As innovation propels and datasets develop, further refinement and investigation of AI procedures vow to extend how we might interpret the IPL environment and lift cricket examination higher than ever.

VI. FUTURE WORK

Looking forward, the domain of IPL investigation involving AI presents various roads for future investigation and refinement. Right off the bat, the joining of greater and various datasets, including constant match measurements, online entertainment feeling examination, and player wellness measurements, could improve the prescient force of models and give a more far-reaching comprehension of player and group elements. Furthermore, the consolidation of cutting-edge AI procedures, for example, profound learning and support learning holds guarantee for catching complex examples and enhancing vital dynamic continuously situations. Besides, utilizing normal language handling (NLP) calculations to investigate player interviews, media inclusion, and fan opinion could offer important experiences into the mental parts of player execution and group elements. Cooperative examination endeavors between information researchers, sports investigators, and cricket specialists can cultivate interdisciplinary development and drive the advancement of IPL examination strategies. At last, by embracing arising advances and interdisciplinary coordinated effort, future work in IPL examination utilizing AI can possibly change cricket investigation, engage IPL partners with significant bits of knowledge, and raise the general onlooker experience.

REFERENCES

[1] Kiran Gawande; Sumit Harale; Simran Pakhare, Predictive Analysis of an IPL Match Using Machine Learning, 2022.
[2] Dhonge; Shraddha Dhole; Nikita Wavre; Mandar Pardakhe; Amit Nagarale; Nikhil, IPL Cricket Score And Winning Prediction Using Machine Learning Techniques,2021
[3] G. Sudhamathy and G. Raja Meenakshi, Prediction on IPL Data Using Machine Learning Techniques, 2020.
[4] Priyanka Kumbhar; Gaurav Patil; Gaurav Gawarguru; Tejas Nirmal; Nagesh Panchling, IPL Score Prediction & Analysis,2020.
[5] Mayank Agarwal, Archana Kumar, IPL First Innings Score Prediction Using Machine Learning Techniques, 2023.
[6] G V Gayathri, Srinivasa Reddy Gajjala, Algorithms,Prediction of IPL Match Score and Winner Using Machine Learning 2020.
[7] T. A. Severini; Analytic methods in sports: Using mathematics and statistics to understand data from baseball, football, basketball, and other sports. Chapman and Hall/CRC, 2014.
[8] Kilari Jaswanth; Kencham Arun; B Deva Deekshith; Kilari Ashwik; Pavithra J, Analyzing and Predicting Outcomes of IPL Cricket Data,2022.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)