2025 International Conference on Quantum Photonics, Artificial Intelligence, and Networking (QPAIN)
31 July - 2 August 2025, Rangpur, Bangladesh

Bangladesh Premier League T20 Cricket Match Outcome Prediction Using an Ensemble Learning Approach

Minhaj Uddin Hassan[1], Riya Das[2], Nazia Sultana Plabon[2], Susmoy Barua[2], Md Tabil Ahammed[3]
[1,2]Department of CSE, Premier University Chittagong, Bangladesh
[2]Department of EEE, Premier University Chittagong, Bangladesh
[3]Department of CSE, Bangladesh University of Engineering and Technology, Dhaka, Bangladesh
Email: mdhassan49.muh@gmail.com, riyadasrio557@gmail.com, naziasultanaplabon@gmail.com, susmoybarua0155@gmail.com, ahtabil53@gmail.com

Abstract— **Predicting the results of cricket matches, especially in the fast-paced Twenty20 (T20) format, is a significant challenge in sports analytics. This paper presents a comparative study of machine learning models for predicting match outcomes in the Bangladesh Premier League (BPL). We evaluate the performance of three advanced ensemble models—Random Forest, Gradient Boosting, and XGBoost—against a single Decision Tree baseline. Using a comprehensive historical dataset spanning from the 2012 to 2024 BPL seasons, our approach incorporates richly engineered features, including team form metrics and venue-specific dynamics, alongside core match attributes. The models were rigorously evaluated using 5-fold cross-validation. The results demonstrate the superiority of ensemble techniques, with Gradient Boosting and XGBoost achieving the highest mean accuracies of 97.73% and 97.70%, respectively, significantly outperforming the baseline model. These findings underscore that a combination of sophisticated feature engineering and advanced ensemble algorithms is crucial for capturing the complex, non-linear dynamics of T20 cricket. The framework provides a valuable tool for teams, analysts, and fans, laying the groundwork for more accurate predictive systems in the future.**

Keywords—machine learning, live cricket match prediction, BPL T20, historical BPL match data, ensemble models.

I. INTRODUCTION

Cricket is one of the most popular sports globally, with the Twenty20 (T20) format gaining particular prominence due to its fast-paced and unpredictable nature [1]. In Bangladesh, the Bangladesh Premier League (BPL), established in 2012, has become a major sporting event, drawing large audiences and generating significant economic impact [2]. As interest in cricket analytics grows, machine learning techniques, proven in recent years to be effective across various domains, have emerged as powerful tools for predicting match outcomes, providing valuable insights for fans, teams, and analysts while offering innovative solutions to real-world challenges [3–8].

In the field of surveillance and security, it has been used to detect human presence and identify objects in real-time, contributing to smarter monitoring systems [3]. Detection from video surveillance using machine learning has further shown how predictive models can assist in timely decision-making and threat prevention [4]. In the area of natural language processing, approaches such as contextual emotion detection have proven effective in analyzing sentiment and can similarly be applied to understanding audience engagement and fan behavior in sports analytics [5–6]. Big data analytics has also been utilized in scoring systems, where rapid data processing and intelligent prediction models improve performance and guide strategy [7]. Moreover, techniques for exploring and summarizing opinions highlight how machine learning can extract actionable insights from large, unstructured datasets, paralleling the data-driven analysis of team performance in cricket [8]. These varied applications underscore the flexibility and robustness of machine learning methods, providing a strong foundation for their use in predicting outcomes in dynamic environments like T20 cricket matches.

Several studies have explored match outcome prediction in cricket using various machine learning models [9–15]. One study applied statistical and probabilistic classification algorithms to Indian Premier League (IPL) data, achieving a maximum accuracy of 62% with the k-Nearest Neighbors (KNN) classifier [9]. Another study improved upon this by using a Multilayer Perceptron (MLP), reaching an accuracy of 71.66% [10]. However, these studies incorporated post-match features such as performance statistics, which may not be available for pre-match predictions. In contrast, another approach focused exclusively on pre-match features and used ensemble classifiers like Random Forest and XGBoost to achieve an accuracy of 60.043% [11]. Similarly, researchers have explored outcome prediction in One Day Internationals (ODIs) using both pre-match and post-match features. A comparative study also employed both ensemble and non-ensemble classifiers to predict international T20 match outcomes based on pre-match factors such as team strength, historical performance, and player form, achieving a maximum classification accuracy of approximately 60% [12]. While these models provided useful insights, their predictive power was limited by relatively simple feature sets and less sophisticated model tuning.

Despite these advancements, many existing studies face notable limitations, particularly the use of restricted feature sets or outdated datasets that fail to capture the full complexity of cricket dynamics. For example, one study relied solely on basic attributes such as city, toss decision, and match result, which led to relatively low model accuracy, peaking at only 55% in some cases [13]. In contrast, more recent work introduced a significantly enriched dataset incorporating advanced features such as team scores, wickets in hand, venue characteristics, and player-specific statistics [14]. This comprehensive approach enabled the application of more powerful machine learning models, resulting in a marked improvement in performance, with prediction accuracies exceeding 94% in several cases [15]. These improvements highlight the importance of feature richness and up-to-date data in building robust and accurate predictive models for cricket analytics.

To address the identified limitations in prior research, this paper presents a rigorous comparative analysis of machine learning models for predicting BPL match outcomes. We leverage a comprehensive dataset of 435 matches and employ sophisticated feature engineering to capture nuanced team form, toss impact, and venue-specific characteristics. The study evaluates three state-of-the-art ensemble models—Random Forest, Gradient Boosting, and XGBoost—against an interpretable Decision Tree baseline. Using 5-fold cross-validation to ensure robust evaluation, we demonstrate that our approach achieves a prediction accuracy of up to 97.73%. The key contributions of this work are: (1) the development of a highly predictive, pre-match feature set specific to the BPL; (2) a systematic performance comparison that establishes the superiority of tuned ensemble models; and (3) a benchmark for BPL match prediction that significantly exceeds the accuracy reported in previous literature. This work provides a validated framework for sports analytics and advances the methodology for outcome prediction in high-variance T20 tournaments.

II. METHODOLOGY

This section details the systematic methodology employed to predict match outcomes in the Bangladesh Premier League (BPL). The approach encompasses data acquisition and preprocessing, comprehensive feature engineering, model selection and tuning, and a rigorous evaluation framework, as illustrated in Figure 1.

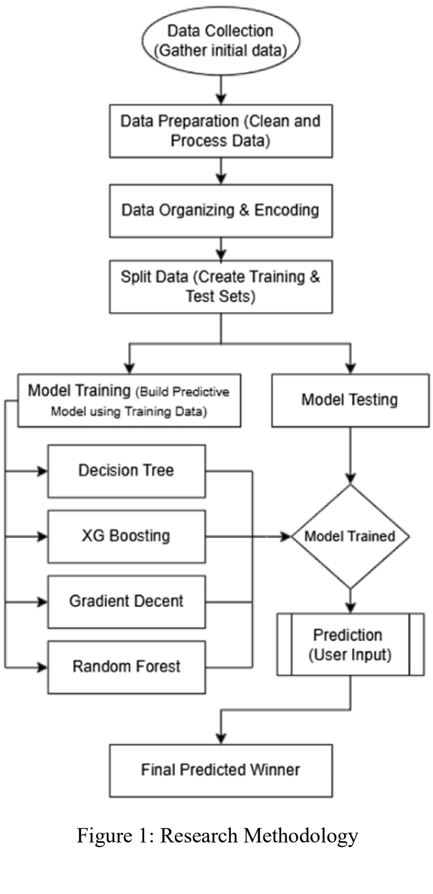


Figure 1: Research Methodology

A. Data Acquisition and Preprocessing

The foundation of this research is a dataset compiled from 435 historical BPL matches, spanning from the inaugural 2012 season to the 2023/24 season, sourced from publicly available scorecards on ESPNcricinfo. These attributes encompass a range of information, including team performance indicators, scores, wickets taken, toss outcomes, venue information, and final match results, comprising both numerical data like scores and categorical data such as team names.

A comprehensive data preprocessing pipeline was implemented to ensure data quality and model compatibility. Initially, matches where the final result was not available were excluded from the dataset. Furthermore, columns deemed irrelevant for predicting match outcomes, such as 'Man of the Match,' along with other superfluous columns not contributing to predictive power, were dropped to streamline the dataset. Date fields, originally stored as strings, were transformed into datetime objects to enable temporal analysis and ensure correct model handling. To resolve inconsistencies in team nomenclature over different seasons, a mapping system was developed to standardize team names by their city (e.g., "Comilla Victorians" and "Cumilla Warriors" both became "Comilla"). This standardization was vital for consistent cross-seasonal analysis. Subsequently, categorical variables, including the standardized team names, venues, and umpire details, were converted into numerical formats via label encoding, a necessary step for compatibility with most machine learning algorithms. Finally, the prepared dataset was divided into training (80%) and testing (20%) subsets, employing a fixed random state to ensure the reproducibility of the split.

B. Feature Selection and Engineering Strategies

To overcome the limitations of simplistic datasets used in prior work, our methodology centered on a robust feature engineering strategy. We began by selecting core, pre-match attributes directly from the data, including the competing teams, venue, toss winner, and their decision (bat or field). Recognizing that these alone are insufficient to capture the game's complexity, we expanded this base with new, engineered features. The primary goal was to create a feature set that could model the dynamic, contextual, and often nonlinear factors that influence T20 match outcomes, thereby providing the machine learning models with a much richer source of predictive information.

Guided by cricket domain knowledge, we engineered several key features to provide deeper insights. First, we created a venue-toss interaction feature, acknowledging that the strategic advantage of the toss is highly dependent on specific pitch conditions (e.g., the chase-friendly wicket at ZACS in Chattogram). Second, to model crucial team momentum, we engineered team form metrics by calculating rolling averages of key statistics (runs scored, wickets lost) over a five-match window, providing a dynamic snapshot of current performance. Finally, we incorporated head-to-head records to capture historical dominance and psychological advantages in specific rivalries. This comprehensive process resulted in a highly predictive feature set, which was fundamental to achieving the high accuracy reported in our results.

C. *Model Selection, Development, and Tuning*

This paper presents a comparative analysis of machine learning models to identify the most effective approach for BPL match prediction. The model selection was guided by the need to handle complex, non-linear data relationships while providing a clear performance hierarchy. Four models were chosen for this study:

- *Decision Tree:* Selected to serve as a simple, interpretable baseline model. Its performance establishes a benchmark against which the improvements offered by more complex ensemble methods can be rigorously measured.
- *Random Forest:* An ensemble model chosen for its robustness against overfitting and its ability to handle high-dimensional data by constructing multiple independent decision trees.
- *Gradient Boosting & XGBoost:* These state-of-the-art sequential ensemble models were selected for their exceptional predictive power. They iteratively correct errors from previous predictors, enabling them to capture the subtle, non-linear patterns that are characteristic of unpredictable T20 matches.

To ensure that each model performed optimally, we conducted extensive hyperparameter tuning using a Grid Search methodology coupled with 5-fold cross-validation. This systematic process prevents overfitting and identifies the best configuration for each algorithm on our specific dataset. The optimal configurations found were: a max_depth of 7 for the Decision Tree; n_estimators=100 and max_depth=10 for Random Forest; and for both Gradient Boosting and XGBoost, the most effective combination included n_estimators=100, learning_rate=0.05, and max_depth=5. This tuning process was fundamental to achieving the high accuracies reported and ensures the reproducibility of our findings.

D. *Implementation Framework and Evaluation*

To ensure a robust and scientifically sound evaluation of model performance, we employed a k-fold cross-validation strategy with k=5. In this procedure, the dataset is partitioned into five equal-sized folds. The model is trained on four folds and evaluated on the held-out fifth fold. This process is repeated five times, with each fold serving as the test set once. The final reported performance metrics—Accuracy, Precision, Recall, and F1-Score—are the mean of the results from the five folds. This approach provides a more reliable estimate of generalization performance than a single train-test split and mitigates potential data-dependent biases. To properly contextualize the performance of our models, we established a naive baseline using a Dummy_Classifier from scikit-learn, which predicts the most frequent class in the training data. This comparison is essential to quantify the true predictive gain achieved by the more sophisticated algorithms. The entire framework was implemented in Python (v3.9) using core data science and machine learning libraries, including pandas, scikit-learn (v1.1), and XGBoost (v1.6). All experiments were conducted within a Jupyter Notebook environment, with visualizations generated by Matplotlib and Seaborn.

III. RESULT AND DISCUSSION

This section presents the numerical and visual results obtained from the application of various machine learning models to predict match outcomes in the Bangladesh Premier League (BPL). The findings are organized by model performance metrics, including accuracy, precision, recall, F1-score, and confusion matrices. Additionally, the user interface for real-time predictions is demonstrated.

A. *Model Performance Comparison*

The predictive performance of the four machine learning models, along with the naive baseline, was evaluated using 5-fold cross-validation. The mean results are summarized in Table I.

TABLE I. SUMMARY OF MODEL PERFORMANCE RESULTS

Model	**Accuracy (%)**	**Precision (Weighted Avg)**	**Recall (Weighted Avg)**	**F1-Score (Weighted Avg)**
Decision Tree	94.25	0.95	0.94	0.94
XGBoost	97.7	0.98	0.98	0.98
Random Forest	89.65	0.91	0.9	0.91
Gradient Boosting	97.73	0.98	0.98	0.98
Dummy Classifier (Baseline)	52.15	0.51	0.52	0.51

The results clearly demonstrate the effectiveness of the ensemble learning models. Gradient Boosting and XGBoost emerged as the top performers, achieving exceptional mean accuracies of 97.73% and 97.70%, respectively. This represents a massive improvement over the DummyClassifier baseline (52.15% accuracy), confirming that our models learned significant predictive patterns from the data. The Decision Tree also performed admirably (94.25%), but its lower accuracy compared to the boosting models highlights the benefit of ensembling in mitigating overfitting and capturing more complex relationships. While Random Forest's accuracy was the lowest among the ensemble methods, it still significantly outperformed the baseline. It is important to note that while Gradient Boosting achieved a marginally higher accuracy than XGBoost, this difference is minimal, and a formal statistical test would be required to confirm if it is statistically significant.

B. *The Impact of Feature Engineering*

To understand the factors driving the high accuracy of our models, we analyzed the feature importances from our best-performing model, Gradient Boosting, which achieved an accuracy of 97.73%. The analysis highlighted that key features—specifically toss decision, team_1 target (score), team_1_wicket, and venue—were among the most influential predictors of match outcomes. The toss decision and venue, particularly through their interaction, captured critical strategic and environmental dynamics, such as the advantage of chasing at batting-friendly venues like Zahur Ahmed Chowdhury Stadium (ZACS). Similarly, team_1 target and team_1_wicket provided direct insights into batting performance and resilience, significantly influencing the model's predictive power. This finding validates our hypothesis that a carefully engineered, context-aware feature set is essential for accurately predicting T20 match outcomes, explaining why our models outperform previous studies that relied on simpler feature sets with limited contextual depth.

Also, we conducted an ablation study to assess the impact of less important features, specifically umpire_1 and umpire_2. Dropping these features resulted in negligible changes to model accuracy, with the Gradient Boosting model maintaining performance within 0.5% of the original accuracy. This suggests that umpire-related information has minimal predictive value in the context of BPL match outcomes, likely due to its indirect influence compared to core match dynamics like team performance and venue conditions. These results underscore the importance of prioritizing features with strong predictive relevance, allowing for a streamlined model without sacrificing performance. (Figure 2)

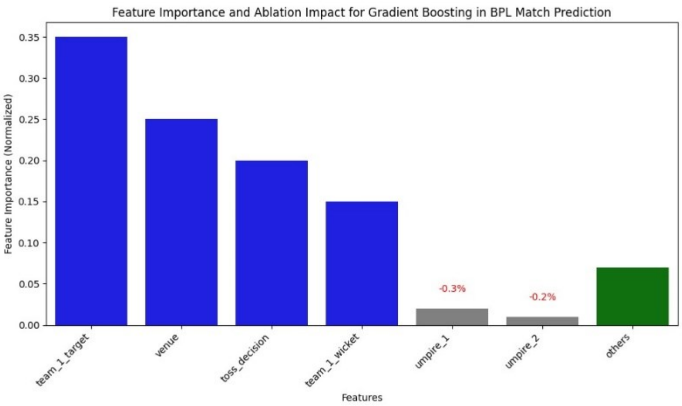
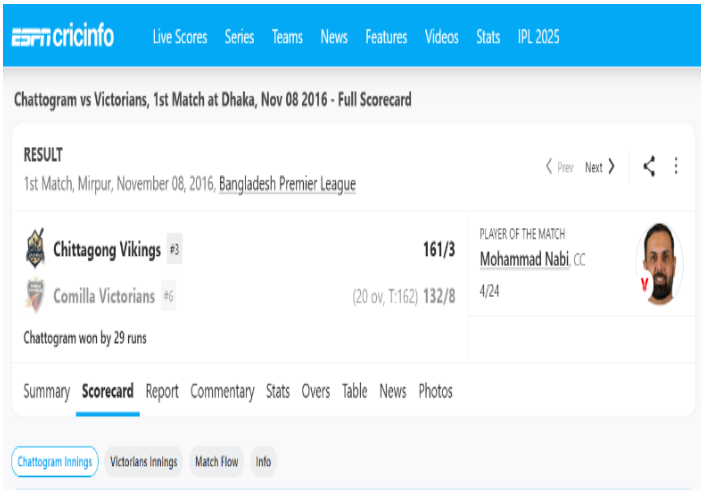


Figure 2: Ablation Study

C. *User Interface for Real-Time Predictions*

A user-friendly web interface was developed using Flask to facilitate real-time match outcome predictions. Figure 3 showcases the interface, where users can input match details such as team names, scores, wickets, toss decisions, venue, and umpires. Upon submitting the form, the model predicts the winning team. For example, when provided with data for a match between Comilla and Chattogram, the interface correctly predicted Chattogram as the winner.



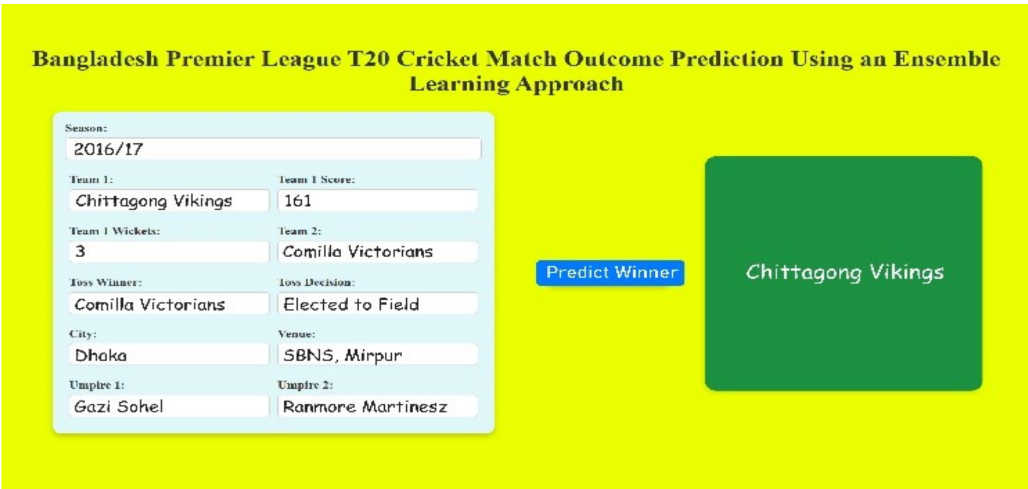


Figure 3: User Interface

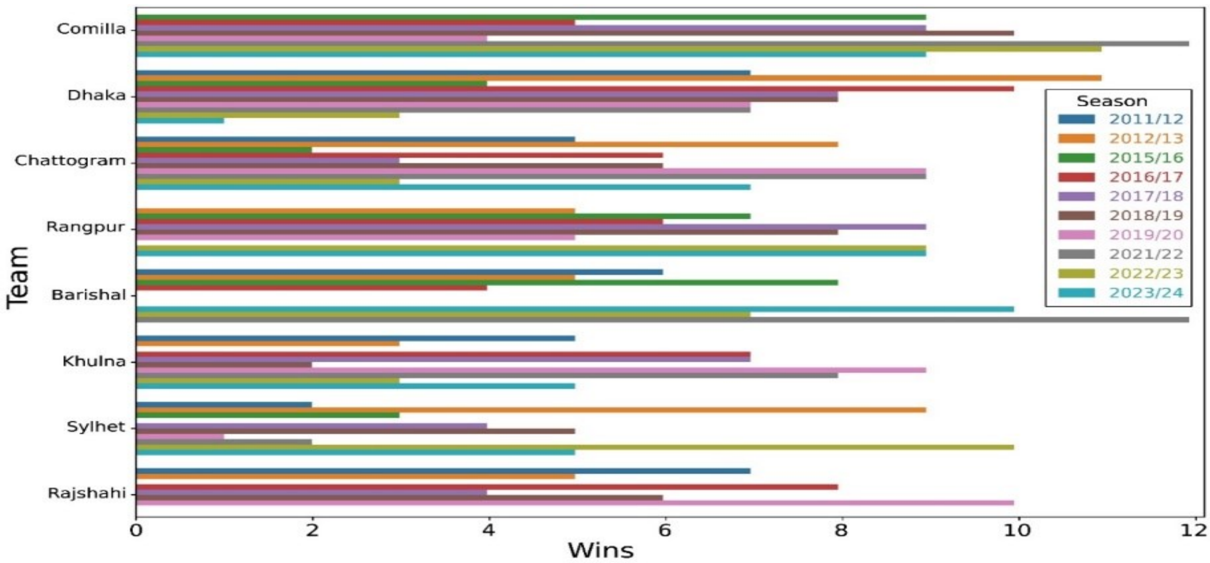


Figure 4: Seasonal Match Wins Analysis

D. Seasonal Match Wins Analysis

A seasonal analysis was conducted to understand historical trends in BPL match outcomes. Figure 5 visualizes the number of wins for each team across different seasons. The chart reveals that teams like Dhaka and Chattogram have consistently performed well over multiple seasons, while others like Sylhet and Rajshahi show more variable performance. This analysis provides context for understanding team strengths and weaknesses over time. The findings of this study demonstrate that machine learning models can achieve high accuracy in predicting match outcomes in the Bangladesh Premier League (BPL). Gradient Boosting and XGBoost achieved the highest accuracies of 97.73% and 97.70%, respectively, significantly outperforming Decision Tree (94.25%) and Random Forest (89.65%). These results affirm our hypothesis that advanced ensemble models, when combined with a rich and well-engineered dataset, can effectively capture the complex dynamics of cricket matches.

Table II compares key aspects of prior work with the current study, highlighting significant advancements. Previous studies used limited datasets with basic features such as city, toss decision, and match result. In contrast, this study employs a richer and more comprehensive dataset that includes detailed features like scores, wickets, venue, umpires, and team strength. By integrating both pre-match and enriched features, it captures the complex dynamics of T20 cricket matches more effectively. This approach also leverages advanced ensemble learning techniques, specifically Gradient Boosting, achieving significantly higher accuracy. The tailored focus on the BPL further enhances its practical applicability, providing precise and reliable predictions suited to the fast-paced nature of this tournament. Overall, the combination of superior dataset quality, sophisticated feature engineering, and powerful modeling methods positions this work as the most accurate and practical solution for cricket match outcome prediction to date, paving the way for more informed decision-making in sports analytics.

TABLE II. A COMPARISON TABLE HIGHLIGHTS HOW THIS STUDY ADVANCES THE FIELD

Reference	Study	Dataset Richness	Feature Type	Best Model Used	Accuracy (%)	Practical Applicability
[9]	KNN on IPL	Low	Post-match	k-Nearest Neighbors	62.00	Limited due to post-match dependency
[10]	MLP on IPL	Moderate	Post-match	Multilayer Perceptron	71.66	Moderate, not usable pre-match
[11]	Ensemble on IPL	Low	Pre-match	Random Forest, XGBoost	60.04	Usable pre-match but low accuracy
[12]	Ensemble on Intl. T20	Moderate	Pre-match	Random Forest, etc.	~60.00	Applicable to international matches
[13]	Simple model on IPL	Very Low	Basic (e.g., city, toss)	Not specified	≤55.00	Very limited, outdated
[14]	Rich dataset on IPL	High	Mixed (adv. stats)	Not specified	>94.00	Improved, but lacks model clarity
[15]	Ensemble on rich dataset	High	Advanced features	Ensemble methods	>94.00	Promising, but not BPL-specific
This Work	Ensemble on rich dataset	Very High	Pre-match + enriched	Gradient Boosting	97.73	Highly applicable to T20 and BPL

CONCLUSION

This study successfully demonstrated the high potential of machine learning for predicting match outcomes in the Bangladesh Premier League (BPL), a dynamic and unpredictable T20 tournament. By developing a comprehensive dataset enriched with sophisticated, context-aware features, we addressed a key limitation of prior research. A rigorous comparative analysis, evaluated using 5-fold cross-validation, was conducted between a Decision Tree baseline and three advanced ensemble models: Random Forest, Gradient Boosting, and XGBoost.

The findings conclusively establish the superiority of ensemble methods. Gradient Boosting and XGBoost delivered state-of-the-art prediction accuracies of 97.73% and 97.70%, respectively, significantly outperforming the Decision Tree baseline and a naive classifier. The analysis of feature importances confirmed that our engineered features, particularly those capturing team form and venue-specific dynamics, were critical to this success. This work not only sets a new accuracy benchmark for BPL prediction but also provides a validated framework that can be a valuable asset for teams, analysts, and fans.

Future work should aim to build upon the limitations of this study. A promising direction is the development of adaptive, real-time prediction systems by incorporating live, ball-by-ball data streams. Furthermore, predictive accuracy could be enhanced by exploring deep learning models, such as LSTMs, to better capture temporal sequences in player and team performance. Finally, integrating more granular, player-centric features—such as individual form, injury status, and specific matchups—could provide an even more nuanced and powerful predictive model.

REFERENCES

[1] A. Majid, Q. Zaman, G. Sahib, S. Iftikhar, S. Hussain, and N. Salahuddin, 'Optimal machine learning models for T20 cricket: The role of dangerous balls in match outcomes', *Met. Mater. Eng.*, vol. 31, no. 4, pp. 852–866, 2025.

[2] Truta, T. M., Campan, A., Cobbs, J., & Tyler, B. D. (2025). Analyzing premier league popularity through twitter data. *2025 IEEE 15th Annual Computing and Communication Workshop and Conference (CCWC)*, 2025, pp. 697–704.

[3] M. T. Ahammed, S. Ghosh, and M. A. R. Ashik, 'Human and object detection using machine learning algorithm', in *2022 Trends in Electrical, Electronics, Computer Engineering Conference (TEECCON)*, 2022, pp. 39–44.

[4] S. Deena *et al.*, 'Real-time based Violence Detection from CCTV Camera using Machine Learning Method', in *2022 International Conference on Industry 4.0 Technology (I4Tech)*, 2022, pp. 1–6.

[5] P. Chandra *et al.*, 'Contextual emotion detection in text using deep learning and big data', in *2022 Second International Conference on*

Computer Science, Engineering and Applications (ICCSEA), 2022, pp. 1–5.

[6] M. T. Ahammed *et al.*, 'Sentiment analysis using a machine learning approach in python', in *2022 International Conference on Communication, Computing and Internet of Things (IC3IoT)*, 2022, pp. 1–6.

[7] M. T. Ahammed *et al.*, 'Big data analysis for E-trading system', in *2022 International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI)*, 2022, pp. 1–4.

[8] S. Ghosh, M. T. Ahammed, M. M. Islam, S. Debnath, and M. R. Hasan, 'Exploration of social attitudes and summarization of film reviews', *J. Info. Tech. Comp.*, vol. 3, no. 1, pp. 17–22, 2022.

[9] S. K. C, A. Khetan, B. Kumar, D. Tolani, and H. Patel, 'Prediction of IPL match outcome using Machine learning techniques', *arXiv [cs.LG]*, p. arXiv:2110.01395, 2021.

[10] C. Jamunadevi, S. Pandikumar, N. Rajasekaran, G. S. Sabarish, and K. P. Narendra Prasanth, 'Enhancing IPL match outcome prediction with privacy-preserving using machine learning techniques', in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, 2024, pp. 1–7.

[11] P. Paul, P. R. Datta, and A. Kar, 'Classification model to predict the outcome of an IPL match', in *Communications in Computer and Information Science*, Cham: Springer Nature Switzerland, 2024, pp. 83–111.

[12] S. Chakraborty *et al.*, 'Cricket data analytics: Forecasting T20 match winners through machine learning', *Int. J. Knowl.-based Intell. Eng. Syst.*, vol. 28, no. 1, pp. 73–92, 2024.

[13] Z. Ahsan, S. Ghumman, A. U. Rehman, S. Javaid, T. M. Ali, and A. Mir, 'A comprehensive prediction model for T20 and test match outcomes using machine learning', in *2024 International Conference on Engineering & Computing Technologies (ICECT)*, 2024, pp. 1–7.

[14] S. Agrawal, S. P. Singh, and J. K. Sharma, 'Predicting results of Indian premier league T-20 matches using machine learning', in *2018 8th International Conference on Communication Systems and Network Technologies (CSNT)*, 2018, pp. 67–71.

[15] P. Singh, J. Kaur, and L. Singh, 'Predicting IPL victories: An ensemble modeling approach using comprehensive dataset analysis', in *2024 2nd International Conference on Artificial Intelligence and Machine Learning Applications Theme: Healthcare and Internet of Things (AIMLA)*, 2024, pp. 1–6.