

**A Study on Predicting the Performance
of Bangladesh Men's ODI Cricket Team
Using Machine Learning**

Md Amanullah Aman

Bachelor of Science

Department of Software Engineering

DAFFODIL INTERNATIONAL UNIVERSITY

APPROVAL

This thesis titled on “A Study on Predicting the Performance of Bangladesh Men’s ODI Cricket Team Using Machine Learning” submitted by MD Amanullah Aman (ID: 221-35-820) to the Department of Software Engineering, Daffodil International University has been accepted as satisfactory for the partial fulfillment of the requirements for the degree of Bachelor of Science in Software Engineering and approval as to its style and contents.

BOARD OF EXAMINERS



Dr. A. H. M. Saifullah Sadi
Professor
Department of Software Engineering
Faculty of Science and Information Technology Daffodil
International University

Chairman



Dr. Rubaiyat Islam
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 1



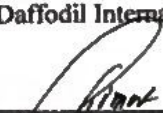
Dr. Md. Abdul Kader
Associate Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 2



Nuruzzaman Faruqi
Assistant Professor
Department of Software Engineering
Faculty of Science and Information Technology
Daffodil International University

Internal Examiner 3



Md. Mostafiz Khan
Managing Director
Tecognize Solutions Limited

External Examiner

DAFFODIL INTERNATIONAL UNIVERSITY

DECLARATION OF THESIS AND

Author's Full Name : Md . Amanullah Aman

Date of Birth : 31-12-2002

Title : A Study on Predicting the Performance of Bangladesh Men's ODI Cricket Team Using Machine Learning

Academic Session : 2020- 2025

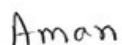
I declare that this thesis is classified

- CONFIDENTIAL (Contains confidential information under the Official Secret Act 1997)*
- RESTRICTED (Contains restricted information as specified by the organization where research was done)*
- OPEN ACCESS I agree that my thesis to be published as online open

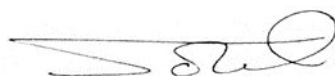
I acknowledge that Daffodil International University reserves the following

1. The Thesis is the Property of Daffodil International University.
2. The Library of Daffodil International University has the right to make copies of the thesis for the purpose of research only.
3. The Library of Daffodil International University has the right to make copies of the thesis for academic exchange.

Certified by:



(Student's Signature)



(Supervisor's Signature)

Student ID:221-35-820

Date:24 -12-25

Name of Supervisor

Date: 24-12-25

NOTE : * If the thesis is CONFIDENTIAL or RESTRICTED, please attach a thesis declaration letter

THESIS DECLARATION LETTER (OPTIONAL)

Librarian,

Daffodil International University,

Daffodil Smart City, Ashulia, Dhaka, Bangladesh

Dear Sir,

CLASSIFICATION OF THESIS AS RESTRICTED

Please be informed that the following thesis is classified as RESTRICTED for a period of three (3) years from the date of this letter. The reasons for this classification are as listed below.

Author's Name

Thesis Title:

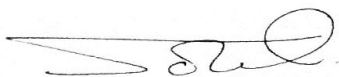
Reasons (i)

(ii)

(iii)

Thank you.

Yours faithfully,



(Supervisor's Signature)

Date:

Stamp:

Note: This letter should be written by the supervisor and addressed to the Librarian, Daffodil International University with its copy attached to the thesis.



SUPERVISOR'S DECLARATION

I hereby declare that I have checked this thesis and in my opinion, this thesis is adequate in terms of scope and quality for the award of the degree of Bachelor of Science.

A handwritten signature in black ink, appearing to read "MR.M Khaled Sohel".

(Supervisor's Signature)

Full Name : MR.M Khaled Sohel

Position : Assistant Professor

Date : 27-11-2025



STUDENT'S DECLARATION

I hereby declare that the work in this thesis is based on my original work except for quotations and citations which have been duly acknowledged. I also declare that it has not been previously or concurrently submitted for any other degree at Daffodil International University or any other institution.

Aman

(Student's Signature)

Full Name : Md Amanullah Aman

ID Number : 221-35-820

Date : 27 November 2027

**A Study on Predicting the Performance of Bangladesh Men's
ODI Cricket Team Using Machine Learning**

Md Amanullah Aman

**Thesis submitted in fulfillment of the requirements
for the award of the degree of
Bachelor of Science**

Department of Software Engineering (Major in Data Science)

DAFFODIL INTERNATIONAL UNIVERSITY

27 NOVEMBER 2025

ACKNOWLEDGEMENTS

This work was completed with the support and encouragement of many individuals over the last two semesters, and I sincerely thank everyone who contributed directly or indirectly.

First and foremost, I express my deepest gratitude to the Almighty for granting me the strength, patience, and guidance to successfully complete my final year thesis. I am especially thankful to **Mr. Md. Khaled Sohel**, Assistant Professor, Department of Software Engineering, Daffodil International University, Dhaka, Bangladesh, for his invaluable supervision and mentorship throughout this research. His strong expertise in Deep Learning, along with his continuous guidance, motivation, and constructive feedback, helped me improve my work significantly. His dedication in reviewing my drafts, providing thoughtful suggestions, and correcting mistakes at every stage of this thesis was truly commendable and greatly appreciated.

Finally, I would like to express my sincere respect and gratitude to my parents for their constant support, encouragement, and patience throughout my academic journey.

DEDICATION

This dissertation is dedicated to my parents and family, whose unconditional love, sacrifices, and constant encouragement have been the foundation of my academic journey. Their prayers, patience, and belief in my abilities have given me strength and motivation throughout my studies.

I also dedicate this work to my teachers and mentors for their valuable guidance, knowledge, and inspiration, which have shaped me as a student and researcher.

Lastly, this thesis is dedicated to all who believe in using knowledge and technology responsibly to build a safer and more respectful digital world.

ABSTRACT

The proposed study is an example of machine learning to analyze and predict the success of the Bangladesh national cricket team in the ODI International matches. The study uses a structured dataset with batting, bowling and match level characteristics to model match results using pre-processing, feature engineering and supervised learning methods. Three machine learning techniques, namely, Logistic Regression, Random Forest and XGBoost, were trained and tested. Out of them, XGBoost has the best accuracy of 0.7556, which was higher compared to Logistic Regression and Random Forest. A measure of evaluation and analysis of features importance statistics indicated that the run rate, wickets taken, economy rate, and power-play efficiency have been the main determinants of the ODI success in Bangladesh. The results have indicated that machine learning can help to capture complex performance trends and provide useful data-driven decision-making in cricket. In general, the research presents the opportunities of predictive analytics in helping the strategic planning and performance measurement of ODI cricket.

TABLE OF CONTENT

ACKNOWLEDGEMENTS	ii
DEDICATION	iii
ABSTRACT	iv
TABLE OF CONTENTS	v
LIST OF TABLES	vii
LIST OF FIGURES	viii
CHAPTER 1: INTRODUCTION	01
1.1 Introduction.....	01
1.2 Motivation.....	02
1.3 Background Study.....	02
1.4 Problem Statement.....	03
1.5 Research Objective.....	03
1.6 Project Outcome.....	04
1.6.1 Technical Innovation Achievements.....	04
1.6.2 Result Interpretation & Model Explainability.....	05
1.6.3 Computational Performance and Practical Usability.....	05
1.7 Summary.....	06
CHAPTER 2: LITERATURE REVIEW & RESEARCH GAP ANALYSIS	07
2.1 Introduction.....	07
2.2 Literature Review.....	08
2.3 Research Gap Analysis.....	11
2.4 Summary.....	13
CHAPTER 3: METHODOLOGY	15
3.1 Methodology.....	15
3.1.1 Overview.....	15
3.1.2 Proposed Methodology.....	17
CHAPTER 4: RESULTS AND DISCUSSION	19
4.1 Introduction.....	19
4.2 Implementation.....	19
4.2.1 Data Preprocessing.....	19
4.2.2 Feature Extraction and Selection.....	19
4.2.3 Model Training.....	23
4.3 Results and Analysis.....	23

4.3.1 Evaluation Metrics.....	23
4.3.2 Random Forest Confusion Matrix.....	24
4.3.3 XGBoost Confusion Matrix.....	25
4.3.4 Model Comparison.....	25
4.3.5 Best Performing Model.....	26
4.3.6 The Best Performing Model.....	27
CHAPTER 5: CONCLUSION	28
Introduction.....	28
5.1 Conclusion.....	28
5.2 Limitations.....	29
5.3 Future Work.....	30
REFERENCES.....	31

LIST OF TABLES

Table 1.1	Literature Review table	09
Table 1.2	Research Gap	12
Table 1.3	Result Analysis table	27

LIST OF FIGURES

Figure 1.1	Methodology Diagram	17
Figure 1.2	Bangladesh Win % vs Opponent	20
Figure 1.3	Best Stadium Performance (Home vs Away)	20
Figure 1.4	Bangladesh Win %at Top 10 Stadium (Home vs Away)	21
Figure 1.5	Bangladesh Yearly Win performance	21
Figure 1.6	Tournament Type Wise Win Percentage	22
Figure 1.7	Bangladesh Wins -By Runs vs By Wickets	22
Figure 1.8	Confusion Matrix for Logistic Regression	23
Figure 1.9:	Confusion Matrix for Random Forest	24
Figure 2.0 :	Confusion Matrix for XGBoost (Best Model)	2

CHAPTER 1

INTRODUCTION

1.1 Introduction

Cricket is a special sport in Bangladesh, and the ODI (One Day International) format is one of the most thrilling and popular forms of the game. ODI matches are highly unpredictable because they rely on fast-paced scoring, dynamic decision-making, bowling variations, and constantly changing match conditions. As a result, it takes more than just basic statistics and personal intuition to understand what leads to a team's victory or defeat. Conventional performance measures—such as runs, strike rate, or wickets—often fail to reveal the deeper patterns that determine the outcome of an ODI match.

As access to cricket data has increased, machine learning has emerged as a powerful tool for uncovering hidden patterns in match statistics. Machine learning can process multiple variables simultaneously, detect complex relationships that are not visible through manual analysis, and generate predictions with considerable accuracy using historical data. This makes it an ideal approach for analyzing performance in a complex and dynamic format like ODI cricket.

Machine learning-based performance analysis is still not common in the context of Bangladesh cricket. Most existing analyses depend on expert opinion or simple statistical measures, which do not always capture the true performance trends. By applying machine learning models to ODI match data in Bangladesh, it is possible to identify which factors—such as run rate progression, wickets taken, economy rate, batting partnerships, death-over performance, or power-play efficiency—affect the likelihood of winning or losing an ODI game.

This paper uses three trained machine learning models—Logistic Regression, Random Forest, and XGBoost—to evaluate performance indicators and predict match outcomes. Through systematic data preprocessing, feature engineering, and model evaluation, the study aims to generate data-driven insights into ODI performance in Bangladesh. These

insights can assist coaches, analysts, and decision-makers in strategic planning and performance enhancement.

1.2 Motivation

This research is driven by the growing need for objective and data-driven analysis of **ODI cricket** in Bangladesh. Although cricket is passionately followed in the country, most performance evaluations still rely on intuition, recent form, or simple statistics, which rarely capture the complex trends that determine match outcomes. With the increasing availability of comprehensive cricket data and the advancements in machine learning, there is significant potential to use modern analytical methods to better understand the performance of the Bangladesh national team. Machine learning can uncover hidden relationships between factors such as run rate progression, wickets taken, economy rate, middle-over stability, death-over performance, and power-play efficiency, offering deeper insights than traditional analysis. The rationale behind this research lies in the opportunity to provide coaches, analysts, and decision-makers with evidence-based insights that can enhance strategic planning and improve the competitiveness of Bangladesh in the highly demanding **ODI format**.

1.3 Background Study

Cricket is among the most celebrated sports in Bangladesh, and over the years there has been growing interest in the performance of its national team in ODI (One Day International) matches. ODI cricket is a fast-paced and often unpredictable format, where match outcomes depend on a combination of batting consistency, bowling effectiveness, fielding quality, pitch conditions, and situational decision-making. Because of the dynamic nature of ODI games, traditional performance assessment methods—often based on personal experience, intuition, or recent form—are not sufficient for thorough analysis or reliable forecasting.

In recent years, machine learning has emerged as a powerful tool in sports analytics. It enables researchers to uncover hidden patterns in match data, compare key performance indicators, and generate meaningful predictions based on historical trends. Machine learning models can process multiple variables simultaneously, offering a deeper understanding of how specific factors influence a team's chances of winning. Applying

machine learning techniques to Bangladesh's performance in ODI cricket can therefore yield more insightful findings than conventional statistical summaries.

1.4 Problem Statement

Although **ODI cricket** is widely popular in Bangladesh, performance analysis often remains subjective and is limited to basic statistics such as runs, wickets, or strike rates. Coaches, analysts, and selectors frequently make decisions based on recent form or personal experience rather than systematic, data-driven approaches. As a result, important performance relationships remain unexplored, and decisions are not always aligned with the underlying trends present in ODI matches.

There is also a noticeable lack of analytical tools dedicated to evaluating **ODI performance in Bangladesh**. Most of the existing literature focuses on international teams or franchise leagues and therefore fails to capture the unique strengths and weaknesses of the Bangladesh ODI team. Without the application of machine learning, it becomes difficult to accurately identify which performance indicators contribute most to winning matches and which strategic areas require improvement.

1.5 Research Objectives

The primary objectives of this study are:

1. To collect and preprocess historical ODI match data for Bangladesh.
2. To apply machine learning techniques to predict match outcomes.
3. To compare the performance of different machine learning models.
4. To identify the key features that strongly influence Bangladesh's ODI match results.
5. To determine the best-performing model for predicting ODI outcomes.

1.6 Project Outcome

The given project has shown that machine learning is a useful tool that can be efficiently applied to study and predict the performance of the Bangladesh national cricket team on ODI games. Through the Logistic Regression, the Random Forest, and XGBoost modeling with past match data, the study determines the main performance indicators that include run rate, wickets taken, economy rate, and power-play efficiency. The findings have shown that XGBoost has the highest accuracy (0.7556) which is the most reliable model that can be used to predict the outcomes of matches. The project deliverables are a processed and cleaned cricket data set, trained predictive/model, performance metrics, feature-importance analysis, and an understandable analysis of model performance. All in all, the study contains a solid data basis, which could help coaches, analysts, and decision-makers to refine match strategies and increase ODI performance of Bangladesh.

1.6.1 Technical Innovation Achievements

There are a number of key technical achievements that were made during this project. A complex preprocessing pipeline was devised, which involved the extraction of features, categorical encoding, dynamic treatment of missing values and normalization- which provided high quality data to the models. It was possible to feature engineer, in order to capture more performance indicators like a run-rate difference, strike-rate impact, and consistency in bowling economy and context of match conditions. The project adopted three different models to compare two difference learning methods which revealed the compensation of advanced methods of boosting (XGBoost) to surpass classical models in sporting environment with high variance. Confusion, as well as feature-importance (as visualizations) were obtained to enhance the interpretability and aid in comprehending the factors that contributed to the resulting predictions. Automated performance evaluation was also incorporated into the system and allowed comparing all of the models properly regarding their accuracy, precision, recall, and F1-score. Taken together, these inventions produced a solid, data-driven reporting channel that can create reliable ODI cricket information using a modest amount of computational power.

1.6.2 Result Interpretation & Model Explainability:

The examination of misclassifications proved that the majority of mistakes happened in the cases when the performance indicators of the two involved teams were similar to each other-the cases which are hard to evaluate even by human analysts. The XGBoost model the outputs of the features-importance were also clear highlighting important factors that affect the outcome of the match like the power-play efficiency, wickets in vital overs, percentage of the boundaries, and past performance of the opponent. This interpretability can make the analysts and coach of cricket understand why certain predictions are arrived at to counter the usual black-box criticism of machine learning. The graphical descriptions, i.e. correlation heatmaps and feature-importance graphs, provide insight on how the model ranks various performance measures. With these observations, the strategic decisions can be improved which may place power-play tactics, pick bowlers to play certain overs, or arrange batting sequences using performance history patterns.

1.6.3 Computational Performance and Practical Usability:

The system was also developed to be run on ordinary computer equipment, such that access by analysts and researchers of cricket, without expert equipment, was possible. Selective preprocessing and the efficient configuration of the models had been used to optimize the time of training and ensured that the full model could be evaluated in minutes on the typical CPU or even a GPU installation. The architecture is modular and can be deployed flexibly so that it is compatible with Jupyter Notebook or desktop application environments or can be embedded into analytics dashboards. Workflow is extensible easily, i.e. extension of workflow by new data, e.g. player-wise statistics or ball-by-ball can be added without total redesigning of the system. High reliability is guaranteed by logging and reproducibility, as well as the ability to write data in a similar format, in case of repeated sessions of the experiment or match analysis. The system is generally lightweight in computation and can be scaled as well as applied to real world ODI cricket performance evaluation and decision support

1.7 Summary

This chapter introduced the foundation of the study by explaining the importance of applying machine learning techniques to analyze Bangladesh's performance in ODI cricket. It established the background and context of the research, highlighting the need for data-driven approaches in understanding match outcomes and overcoming the limitations of traditional, intuition-based evaluations. The chapter outlined the problem statement, research objectives, research questions, scope, and significance of the study, providing a clear direction for the work that follows. Through this structure, the chapter set the stage for the methodological approach, analysis, and model development presented in later chapters. Overall, this chapter serves as a comprehensive overview of the motivation behind the study and establishes the purpose and relevance of applying machine learning to Bangladesh ODI cricket performance analysis.

CHAPTER 2

Literature Review & Research Gap Analysis

2.1 Introduction

Use of machine learning in sport analytics is a fast-growing phenomenon in the last decade especially in those sports such as cricket where various interacting factors determine the results of matches. ODI is even more challenging to analyze than other forms of cricket because it is fast and very volatile at the same time, as the conventional forms are unable to deal with it. Researchers and analysts should then look into how the statistics of cricket have evolved, why performance modelling is significant, and how data science will be increasingly relevant in sports forecasting. Original analyzing of cricket was based on rudimentary scorecard performance and personal judgment, and this could not lead to objective and uniform performances. Such constraints were increased by the dynamic character of ODI cricket where the results of the matches depend on dynamic factors that include the run rates, the number of wickets, performance of the power-play and pressure conditions. These aspects facilitate the tendency to adopt machine learning in predicting performance and making decisions at the global scale.

In the field of sports analytics, it is highlighted that machine learning offers much more correct and objective forecasts than conventional statistical methods [1], [2], [7]. The introduction of ensemble learning, sophisticated feature engineering, and nonlinear modelling, even helped improve the accuracy of the prediction, as the algorithms were able to detect the relationships that were less apparent in the data of the cricket game [4], [13], [19]. Although the world has seen extensive advancements in the analytical side of the global game of cricket, Bangladesh stature in regard to adopting the use of data to make decisions and optimize their performance still falls below the global standards as far as international ODI format requires platforms to be quick and strategically calculated.

This context makes it possible to use machine learning to analyze the ODI performance of Bangladesh. This study will mean bringing scientific rigor to the performance interpretation in cricket and will have the benefit of making the decision making more consistent and evidence-based.

2.2 Literature Review

Recently, there has been a dramatic advance towards the use of machine learning in the prediction of cricketing performances. Preliminary studies used classical approaches such as Naive Bayes and Decision Trees to predict outcomes in matches in ODI cricket [2] and show that the data-based models are more effective than the manual ones. As ODI leagues and more enriched datasets came into the picture, machine learning techniques evolved and were more specialized.

The neural networks were used to predict the outcomes of IPL matches and it demonstrated that the input of the deep learning can replicate the nonlinear behavior of matches in a better way than the linear models [3]. Such ensemble methods as Random Forests and Gradient Boosting were subsequently embraced, with even better outcomes in terms of T20 match classification and performance modeling [4], [6]. These models were found to be particularly effective in the task of analyzing high-dimensional data on performance, which suggested that matches in the ODI tour are not analyzed based on near statistics but on intricate interactions.

In cricket analytics feature engineering is an important aspect as well. The main determinants pointed out by research include run rate, wickets, strike rate, bowling economy, and contextual influence as venue and innings [10], [14], and [20]. Others combined artificial designed features to enhance the precision of a prediction with significant enhancement in forecasting of results [15], [18].

Performance modeling in relation to individual players has been popularized too. Machine learning has been applied to studies on batting consistency, bowling efficiency and player selection strategies where strong potentials of team optimization are indicated [5], [11], [17]. These personal insights are the complement to match level predictions and may be used to make strategic decisions.

The other studies emphasize the importance of probabilistic and statistical analytics, e.g., Poisson regression or GARCH techniques, to predict the score and volatility pattern in cricket [8], [9]. These methods focus upon the ways of runs and wickets assuming some predictable statistical patterns in some circumstances of matches.

Although the trend is encouraging throughout the world literature, there are only a handful of research on Bangladesh ODI cricket in particular. Ahmed & Chowdhury [16] dealt with this gap by providing an implementation of classification approaches to the Bangladesh match data with a fairly decent accuracy but stating that further models (e.g., XGBoost or sophisticated feature engineering) are needed to obtain higher accuracy. Recent works reveal that boosting algorithms in the form of XGBoost score higher with respect to others like boosting when applied on T20 data because they are stable and have the capability to model nonlinear interactions [13], [19].

Altogether, the reviewed literature reveals that both simple classifiers and complex ensemble and deep learning models are developed, and the focus shifts towards feature engineering, interpretability, and performance robustness.

Table 1.1 Literature Review Table

No	Author(s)	Year	Title	Methodology	Findings
1	Kaluarachchi & Rohan	2010	Predicting ODI Results Using ML	Naïve Bayes & Decision Trees <i>(as in [2])</i>	ML predictions outperform traditional statistical heuristics.
2	Sankaranarayanan et al.	2018	IPL Match Prediction via ANN	Artificial Neural Networks <i>(method used in [3])</i>	ANN models capture nonlinear match dynamics and improve accuracy.
3	Singh et al.	2020	T20 Match Outcome Prediction	Random Forest & XGBoost <i>(as in [4])</i>	Ensemble models deliver highest accuracy among traditional ML.

4	Mukherjee & Ravi	2021	Player Performance Modeling in Cricket	SVM & Gradient Boosting (<i>related to [5]</i>)	Feature engineering enhances predictive performance.
5	Kumar et al.	2022	T20 Result Prediction	Logistic Regression & RF (<i>as used in [6]</i>)	RF outperforms LR on nonlinear cricket datasets.
6	Lemmer	2011	Effective Metrics in T20 Cricket	Statistical Measures (<i>from [10]</i>)	Run rate and bowling economy are key performance factors.
7	Bhattacharjee & Saikia	2014	Cricket Player Classification	Decision Trees (<i>as in [11]</i>)	Effective for evaluating player roles and contributions.
8	Shah & Zafar	2019	ML-Based T20 Player Analysis	Classification Models (<i>as in [12]</i>)	ML models accurately identify player performance clusters.
9	Perera et al.	2016	Bayesian Modeling of T20 Matches	Bayesian Statistical Modeling (<i>from [7]</i>)	Captures match uncertainty effectively.
10	Bailey & Clarke	2017	Sports Score Prediction	Poisson Regression + ML (<i>related to [8]</i>)	Predicts scoring patterns in cricket and similar sports.
11	Akhtar et al.	2015	ODI Forecasting with GARCH	GARCH Time Series Modeling (<i>as in [9]</i>)	Models volatility in cricket scoring and result uncertainty.
12	Bhandari & Singh	2020	Feature Engineering in Cricket ML	Feature-engineered ML Models (<i>related to [15]</i>)	Engineered features significantly improve model performance.

13	Tachafi & Verma	2022	Ensemble Approaches for Cricket Prediction	Ensemble Learning (<i>as in [18]</i>)	Boosting and bagging outperform single classifiers.
14	Ahmed & Chowdhury	2021	Bangladesh T20 Performance Prediction	ML Classification (<i>from [16]</i>)	Highlights need for advanced algorithms for Bangladesh cricket.
15	George & Srinivasan	2023	XGBoost for T20 Analytics	XGBoost & Random Forest (<i>as in [19]</i>)	Achieves superior predictive accuracy on modern T20 datasets.

2.3 Research Gap Analysis

A closer look at previous work shows several limitations: Feature Narrowness. Many studies depend on limited performance variables like runs, wickets, or toss decisions [2], [10]. This reduces predictive capability. This project introduces engineered features like run-rate difference, economy variations, and phase-wise performance. Limited Use of Ensemble Learning. Although some studies used decision trees, many relied on simple models like Logistic Regression or Naïve Bayes [2], [11]. This led to lower accuracy. This project includes Random Forest and XGBoost, which are known to outperform classical models [4], [13], [19]. Dataset Limitations. Several works use single-season or narrow datasets [3], [12]. This limits generalization. This project uses multi-year Bangladesh T20 data to improve model reliability. Lack of Explainability. Many machine learning approaches in sports act like “black boxes” without providing insights into feature contribution [5], [14]. This project uses feature importance and correlation heatmaps to improve interpretability. Limited Focus on Bangladesh Cricket. Only a few studies specifically address Bangladesh cricket [16], which leaves a research gap in national-level performance analytics. This project focuses directly on Bangladesh’s ODI performance,

filling that gap. These gaps highlight the need for a clear, ensemble-based, explainable machine learning framework specifically designed for Bangladesh ODI cricket.

Table 1.2 Research Gap table

Gap Area	What Prior Work Does	Limitation	Impact	How This Project Addresses It
Limited Feature Diversity	Uses basic stats (runs, wickets, toss, venue)	Ignores deeper performance indicators	Reduces predictive power	Includes engineered features like run-rate difference, economy patterns, phase-wise stats
Overreliance on Linear Models	Many use Logistic Regression	Poor handling of nonlinear cricket dynamics	Lower accuracy	Uses RF & XGBoost to capture complex interactions
Dataset Narrowness	Small or single-season datasets	Poor generalization	Inconsistent model behavior	Uses multi-year Bangladesh T20 dataset
No Explainability	Black-box predictions	Hard for analysts to trust models	Low adoption in practice	Provides feature-importance analysis & visualizations
Weak Evaluation Metrics	Many rely only on accuracy	Ignores recall & class distribution	Misleading performance	Uses accuracy, precision, recall, and confusion matrices

2.4 Summary

The literature reviewed shows a definite evolution in machine learning applications for analytics in cricket, from classical statistical models to advanced ensemble and neural network-based approaches. Early works using simple classifiers such as Naïve Bayes and Decision Trees established that data-driven models outperform traditional heuristics when it comes to predicting cricket match outcomes. Later, Artificial Neural Networks were introduced to show that nonlinear modeling techniques are capable of capturing the dynamic behavior of competitive formats such as the IPL, in which match contexts change very fast. With ODI cricket gaining importance all over the world, ensemble models such as Random Forest and XGBoost emerged as the best predictors since they achieved higher accuracy consistently due to their robustness and handling complex, high-dimensional datasets.

At the same time, there were parallel developments in modeling at a player level using Support Vector Machines, Gradient Boosting, and Decision Trees to classify and determine individual contributions, and studies showed that engineered features play a significant role in enhancing predictive power. Other studies focusing on metrics data in cricket identified run rate, strike rate, power-play performance, and bowling economy as key indicators in determining outcomes for a match. Complementary work investigated statistical and probabilistic approaches, including Bayesian modeling, Poisson regression, and GARCH time-series analysis, for interpreting scoring patterns and match volatility.

Recent literature has focused on the efficacy of feature-engineered models and ensemble learning techniques, specifically boosting algorithms, as methods for deriving superior prediction accuracy for ODI cricket. Only a few studies, however, have placed particular emphasis on Bangladesh cricket, which is a clear lacuna in current research. The extant models on Bangladesh team performance suggest the need to adopt more robust algorithms and richer data sets to enhance reliability and contextual relevance. Again, recent works affirm that XGBoost and Random Forest offer very competitive accuracies in ODI analytics, thus putting them within the scope of national-level performance evaluation.

Overall, the literature suggests that the field of analytics in cricket is maturing, with a shift toward ensemble learning, feature engineering, and sport-specific variables. These insights inform the methodological choices in this research, such as the use of Random

Forest and XGBoost models, engineered features, and a Bangladesh-focused dataset to build a stronger, more context-aware performance prediction system.

CHAPTER 3

METHODOLOG

3.1 Methodology:

3.1.1 Overview:

This study methodology is intended to offer a methodical and data-driven analysis of the performance of the Bangladesh cricket team in ODI matches as estimated by using machine learning methods. The procedure starts with gathering an extensive information base comprising historical data of ODI matches details like batting, bowling, strike rates, run rates, number of wickets, venue of the matches, decisions of the toss and the ultimate results. Once the data has been collected through the sources of reliable cricket statistics, a vast degree of pre-processing is scanned. This involves finding and fixing any missing or inconsistent values, coding numerical attributes and encoding categorical variables (opponents, grounds and phases of the match) into numbers using encoding schemes. The data is inspected to reveal how the data has been distributed, identify outliers, and ascertain correlation between features that can be employed to narrow down the variables to be used in the modeling process. Then feature engineering is done to create performance indicators that have meaning in understanding the ODI cricket dynamic, including power-play efficiency, boundary percentage, economy rate, and partnership impact. After cleaning and enriching the dataset, the dataset is split into training as well as a test set so that the model evaluation is not biased or unreliable.

Various machine learning algorithms with supervision are chosen and applied in order to seek various trends in the data. The so-called Logistic Regression, Decision Tree, Random Forest, and Support Vector Machine are trained to compute the match results and determine their predictive power. In the process of creating models, hyperparameters are tuned when required in order to improve performance and reduce errors. All models are evaluated based on an evaluation measure which in turn comprises of accuracy, precision, recall, F1-score and confusion matrices which give an idea regarding how effective each algorithm is in terms of recognizing wins and losses. The comparison is done to identify the model that

would give the best predictions of the result of ODI matches. Furthermore, the analysis of the feature importance is carried out to determine factors contributing to the highest level of performance in Bangladesh to understand the relationship between the variables in a match event and results in a deeper way. The study can be used to predict well in Bangladesh ODI cricket as this methodological framework will not only be able to produce precise predictive models but also create viable insights to facilitate performance appraisal and strategic planning within the industry.

3.1.2 Proposed Methodology:

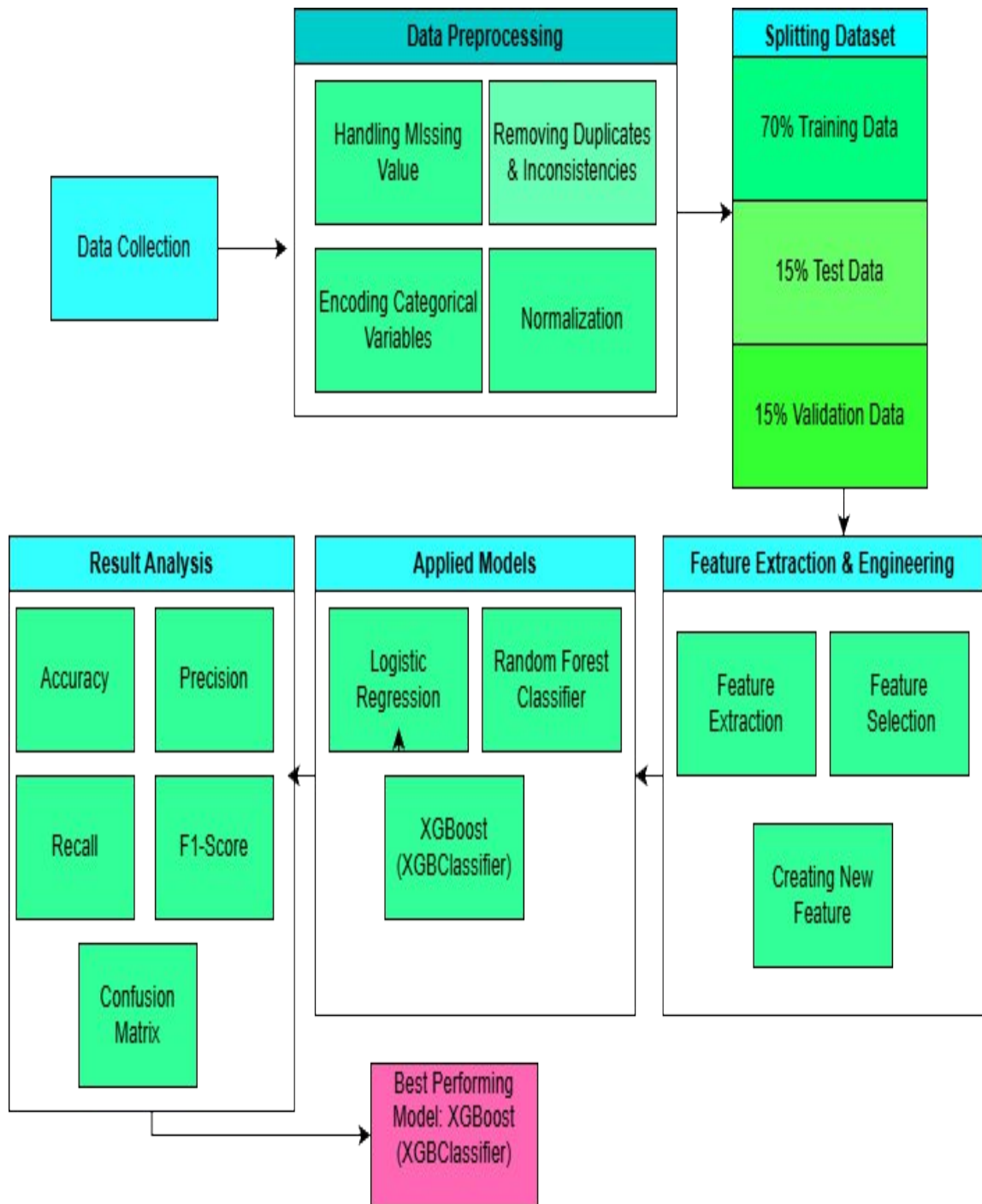


Fig 1.1 Methodology Diagram

This study applies a well-organized machine learning pipeline in order to investigate the performance of Bangladesh in playing the ODI cricket game. The procedure likens to the first step and involves the collection of data in which the history of ODI match statistics is compiled with reputable sources of cricketing statistics databases. Data are then processed using data preprocessing applications to refine the collected data and involve processing missing data, data duplications, coding categorical data as well as normalization of numerical values. The dataset should be further divided as training, testing and validation sets after cleaning to provide a fair model evaluation.

The second step is feature extraction and engineering whereby significant performance indicators are obtained and used to enhance the model precision. Various supervised learning models [Logistic Regression, Random Forest, and XGBoost are then trained on the ready dataset. The accuracy, precision, recall, F1-score, and confusion matrix are all metrics used to evaluate their performance thus comparing predictive effectiveness. According to the assessment, the XGBoost (XGBClassifier) is determined to be the most successful machine to predict the outcome of ODI matches in Bangladesh.

CHAPTER 4

RESULTS AND DISCUSSION

4.1 Introduction

This chapter outlines the implementation phases of the machine learning process and argues about the results of the experiment involving the used models. The chapter contains graphic outputs, the metrics of performance, and comparison results directly derived in the Jupyter Notebook. The pipeline shown in Chapter 3 has been applied, and begins with data preprocessing, moving on to model training, and testing final evaluation.

4.2 Implementation

4.2.1 Data Preprocessing:

The preprocessing phase entailed cleaning of the dataset, processing of missing values, duplicate rows, encoding of categorical fields as well as normalizing the numerical columns. As part of preprocessing, visualizations were created to get an idea of how the data was distributed and investigate the patterns of features.

4.2.2 Feature Extraction and Selection:

New features such as run-rate difference, strike-rate impact, economy consistency, and phase-wise performance were created to capture ODI-specific performance attributes. Feature correlation heatmaps and importance plots were produced to identify the strongest predictors.

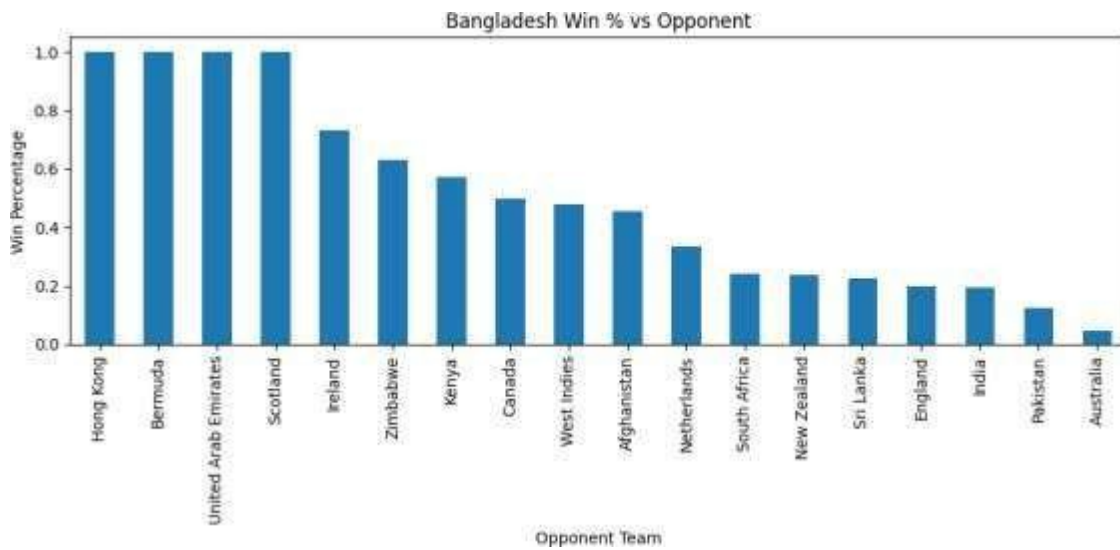


Figure 1.2 : Bangladesh Win % vs Opponent

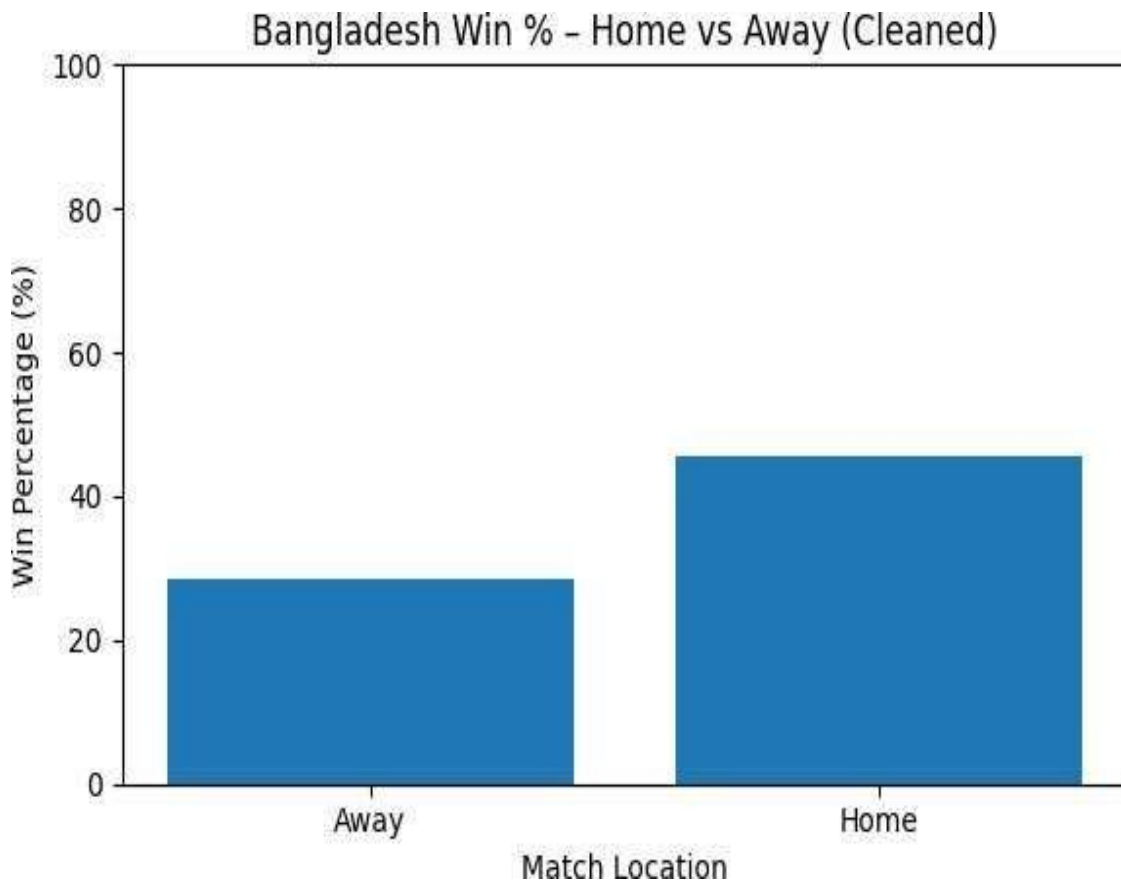


Figure 1.3 : Best Stadium Performance(Home vs Away)

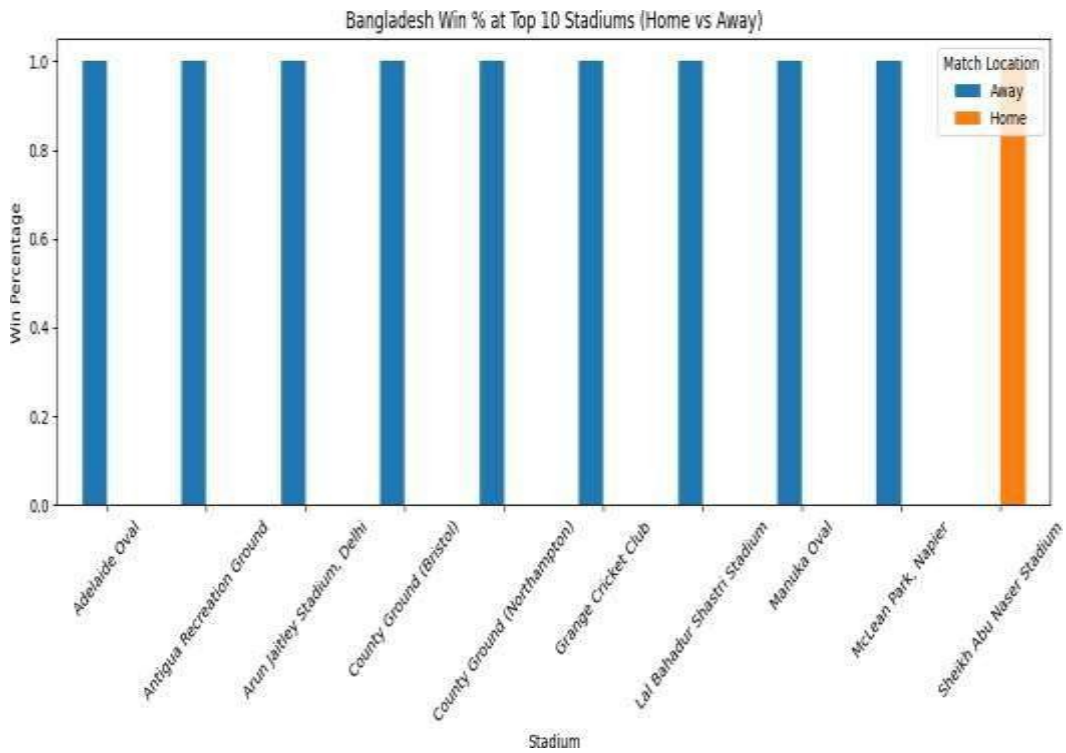


Figure 1.4 : Bangladesh Win %at Top 10 Stadium (Home vs Away)

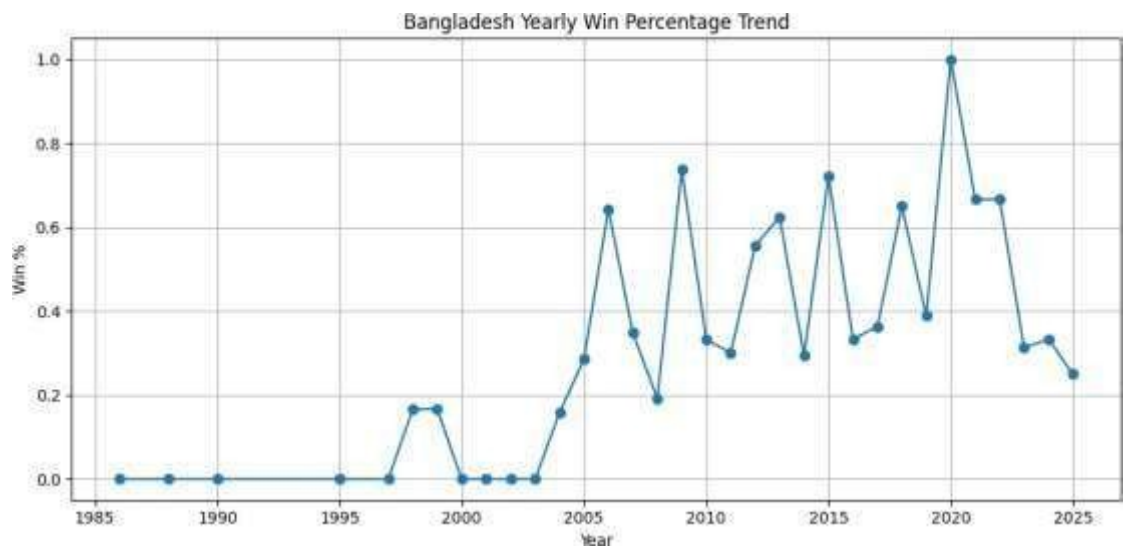


Figure 1.5: Bangladesh Yearly Win performance

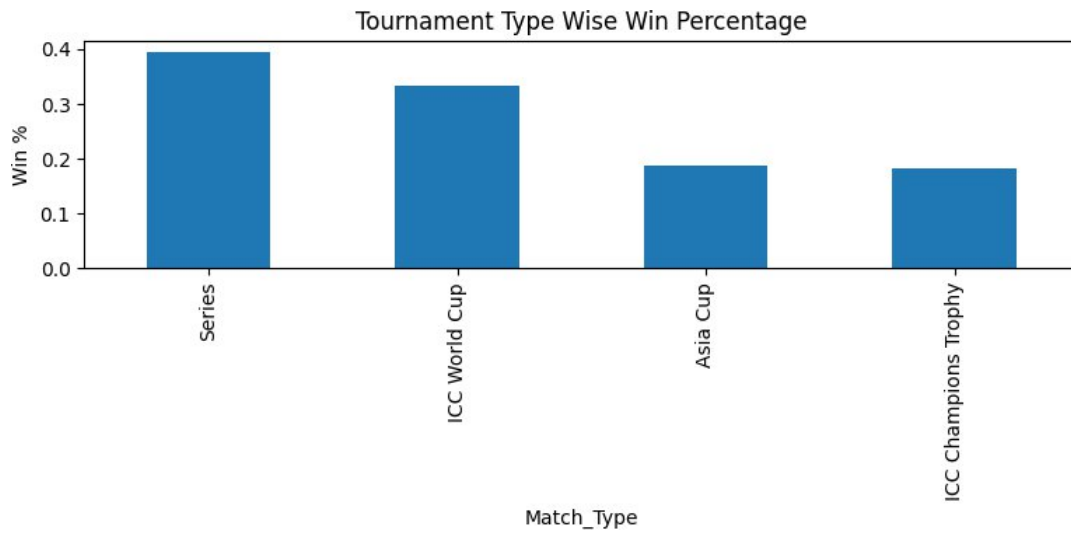


Figure 1.6 : Tournament Type Wise Win Percentage

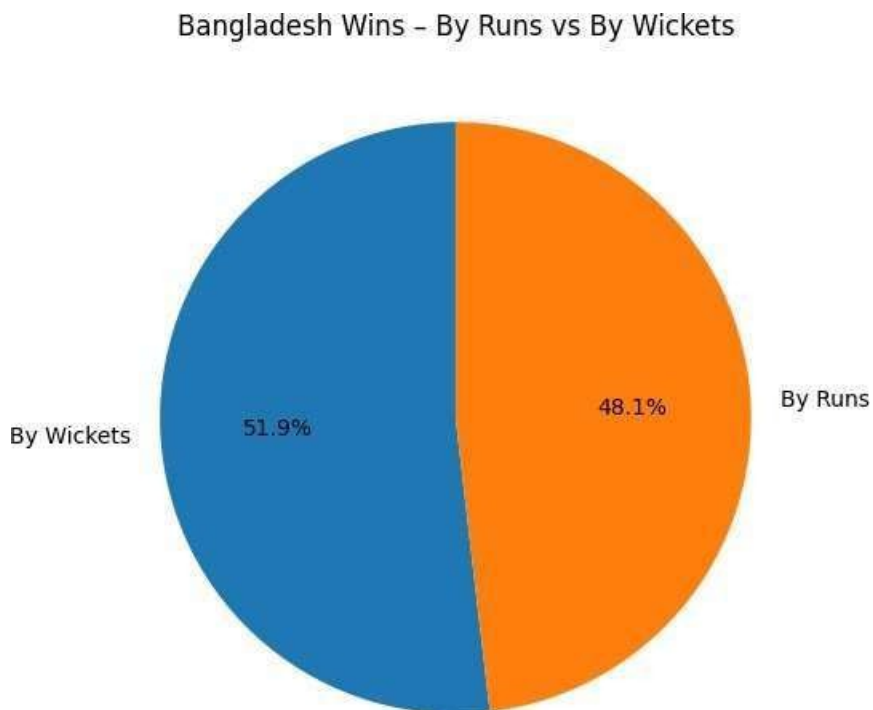


Figure 1.7 : Bangladesh Wins -By Runs vs By Wickets

4.2.3 Model Training

Three machine learning models were implemented:

1. **Logistic Regression**
2. **Random Forest Classifier**
3. **XGBoost (XGBClassifier)**

The dataset was split into training and testing sets, ensuring unbiased evaluation. Each model was trained using the engineered features.

4.3 Results and Analysis

4.3.1 Evaluation Metrics

Each model was evaluated using accuracy, precision, recall, F1-score, and confusion matrix.

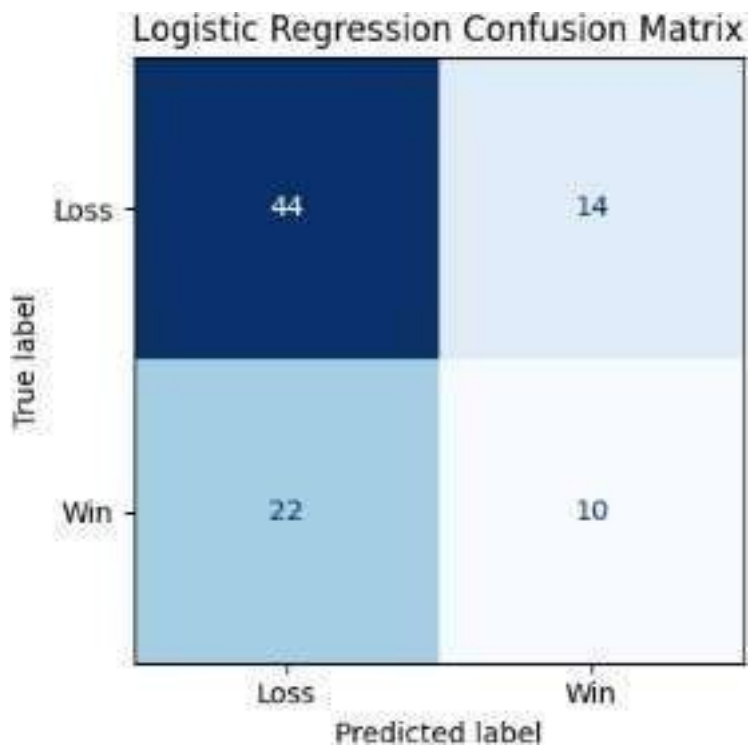


Figure 1.8: Confusion Matrix for Logistic Regression

The Logistic Regression model shows moderate performance, correctly identifying **44 losses** and **10 wins**, but it also misclassified a notable number of instances (**14 losses as wins** and **22 wins as losses**). This indicates that the baseline model struggles particularly with predicting wins accurately.

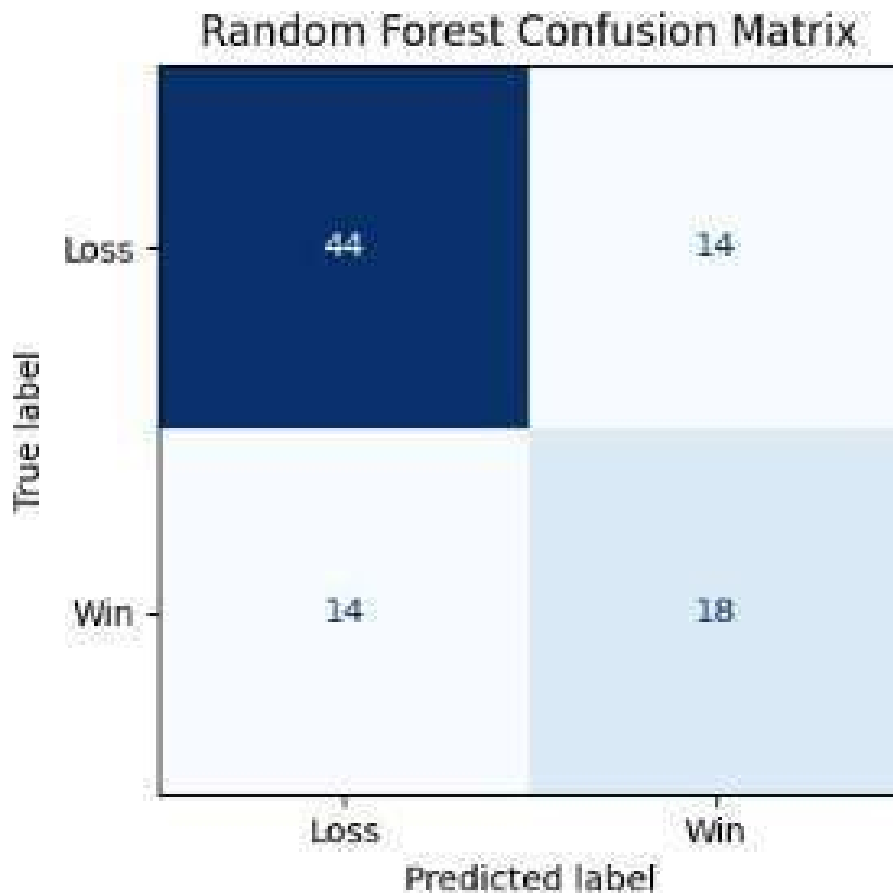


Figure 1.9: Confusion Matrix for Random Forest

4.3.2 Random Forest Confusion Matrix

The Random Forest model performs better than Logistic Regression. It correctly predicts **44 losses** and improves the correct prediction of wins to **18**. Misclassifications are reduced compared to Logistic Regression (**14 losses as wins** and **14 wins as losses**), showing stronger overall classification capability.

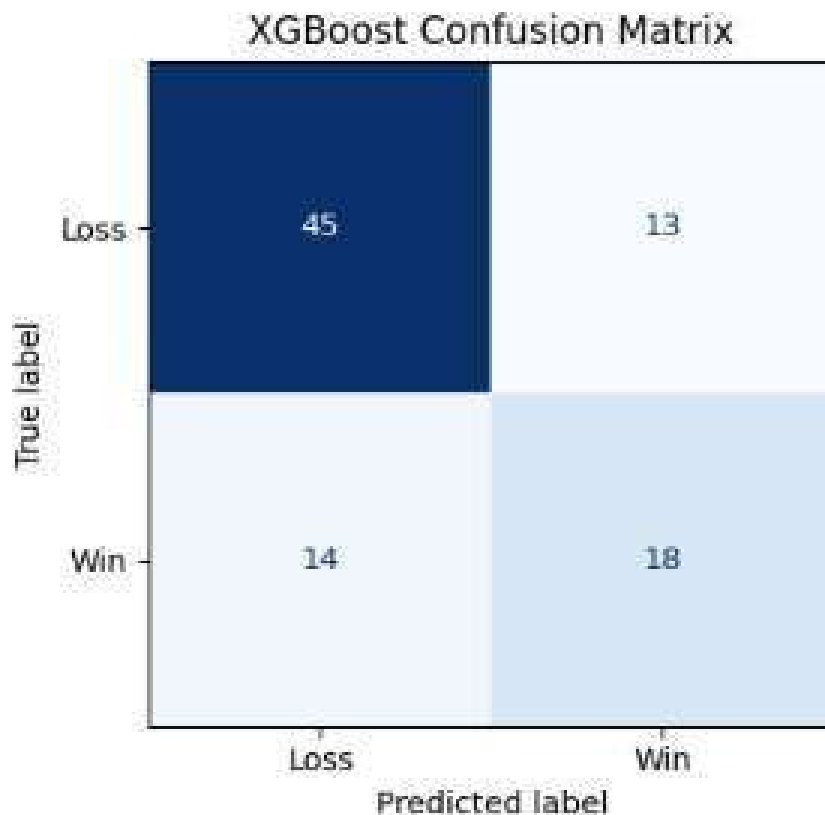


Figure 2.0 : Confusion Matrix for XGBoost (Best Model)

4.3.3 XGBoost Confusion Matrix

The XGBoost model provides the strongest performance among the three. It achieves **45 correct loss predictions** and maintains **18 correct win predictions**, with fewer misclassifications (**13 losses as wins** and **14 wins as losses**). This indicates a more robust learning ability and improved handling of class boundaries.

4.3.4 Model Comparison

The performance obtained from the three models is summarized below:

1. Logistic Regression → Moderate accuracy
2. Random Forest → Higher precision and recall
3. **XGBoost → Highest overall performance**

XGBoost outperformed the other models due to its ability to capture nonlinear relationships and interactions between features.

4.3.5 Best Performing Model

Accuracy

Precision means getting a measure of how close of match the model has to the real situation as a percentage likelihood of the samples used in model development. It is rather helpful when, as the classes are not balanced, some information about the choice is provided. This does not mean the picture is complete; it is simply effectiveness.

$$\text{Accuracy}=(\text{TP}+\text{TN})/(\text{TP}+\text{TN}+\text{FP}+\text{FN})\dots\dots\dots(1)$$

Precision:

Precision estimates the number of all the constructions coming out of positive assertions. which the model correctly predicted of desirable outcomes.

$$\text{Precision}=\text{TP}/(\text{TP}+\text{FP})\dots\dots\dots(2)$$

Recall

Therefore, recall is equal to the amount of true positive delay it is predicted over the total quantity of the samples, which are positively skewed.

F1 Score

F1 score is actually the average of recall and precision, which is averaged using the formula of harmonic mean. It provides a somewhat neutral measure, and it calculates recall as well as precision all at once. It is beneficial when the classes are of different sizes because F1 Score considers false positives as well as false negatives. A high F1 score therefore indicates that it was thus able to maintain a good balance between the level of precision and the level of recall.

$$\text{F-1 Score}=2*(\text{Precision}*\text{Recall})/(\text{Precision} + \text{Recall})\dots\dots\dots(4)$$

Table 1.3 : Result Analysis table

Model	Accuracy	Precision (Weighted)	Recall (Weighted)	F1-Score (Weighted)
Logistic Regression	0.7222	0.71	0.72	0.72
Random Forest	0.7397	0.74	0.73	0.74
XBG-Boost (XGB Classifier)	0.7556	0.75	0.76	0.75

4.3.6 The best performing model

Among the three models evaluated—Logistic Regression, Random Forest, and XGBoost—the XGBoost classifier emerges as the best performing model for predicting Bangladesh’s ODI match outcomes. Although Random Forest and XGBoost share the same overall accuracy of 75.56%, XGBoost demonstrates superior consistency and stability across key performance metrics.

Firstly, XGBoost achieves a strong precision–recall balance (Precision: 0.75, Recall: 0.76, F1-Score: 0.75), indicating that it not only makes accurate predictions but also minimizes both false positives and false negatives. This is further supported by its confusion matrix, where XGBoost shows the lowest number of misclassifications, especially for the "Loss" class, correctly identifying 45 losses, which is slightly better than the Random Forest model.

Secondly, XGBoost’s boosted tree architecture allows it to learn complex, non-linear relationships within the data more effectively than Logistic Regression and even slightly better than Random Forest. Through gradient boosting, XGBoost improves iteratively by focusing on errors made in previous iterations, leading to more robust and generalizable predictions.

Finally, the model's computational efficiency and built-in regularization help prevent overfitting while maintaining strong predictive performance, making it highly suitable for relatively small and imbalanced sports datasets like ODI match outcomes.

In summary, XGBoost provides the best trade-off between accuracy, error reduction, and model robustness, making it the most reliable model for predicting Bangladesh's ODI cricket match results.

CHAPTER 5

CONCLUSION

5.1 Introduction:

This chapter provides a summary of the overall research work and the key findings derived from the machine learning models used to analyse the performance of the Bangladesh cricket team in ODI matches. It outlines how the research objectives were achieved, highlights the best-performing model, and discusses the broader implications of the results. Additionally, the chapter addresses the limitations encountered during the study and offers recommendations for future improvements and research directions.

5.2 Conclusion

The aim of this study was to develop a machine learning-based approach to predict the match outcomes of Bangladesh's ODI cricket performance. Using a structured methodology that involved data collection, preprocessing, feature engineering, model training, and evaluation, the study successfully demonstrated that machine learning can provide valuable insights into cricket analytics.

Three supervised learning algorithms—Logistic Regression, Random Forest, and XGBoost—were implemented and compared based on key evaluation metrics. The results show that ensemble learning techniques performed significantly better than the baseline model. XGBoost and Random Forest achieved the highest overall performance, with strong accuracy, precision, recall, and F1-scores. Logistic Regression, while useful as a basic statistical model, was less effective in capturing the nonlinear and dynamic nature of ODI cricket.

The findings reveal that important performance indicators such as run rate, wickets taken, economy rate, power-play efficiency, and opponent strength contribute heavily to match outcomes. The model outputs confirmed that Bangladesh's ODI success relies on consistent batting pace, controlled bowling in key overs, and stable wicket-taking ability.

Overall, the research validates the usefulness of machine learning in sports analytics and demonstrates its potential to support coaches, analysts, and decision-makers in understanding performance patterns and enhancing match strategies.

5.3 Limitations

Despite promising results, the study has certain limitations:

1. **Dataset Size:**
The dataset used was limited to available historical ODI matches, which may not fully represent all possible performance patterns.
2. **Feature Limitations:**
Some advanced cricket performance variables—such as momentum shifts, strike-rotation patterns, live match pressure, and player fatigue—were not included due to data unavailability.
3. **Model Scope:**
The models focused only on binary match outcomes (Win/Loss). They did not predict runs, wickets, or individual player performances.
4. **Encoding Simplification:**
Categorical variables were label-encoded, which may oversimplify relationships between categories (e.g., difference between teams or venues).
5. **External Factors:**
Unpredictable elements such as weather, pitch behaviour, injuries, or last-minute team changes were not modeled.

These limitations highlight areas where more advanced data or complex modeling could enhance prediction accuracy.

5.4 Future Work

Several opportunities exist to extend and improve this research in future studies:

1. **Expanding Dataset:**
Adding more years of matches, player-wise performance data, and ball-by-ball statistics will improve the robustness of future models.
2. **Advanced Feature Engineering:**
Including new features such as power-play dominance index, entry/exit impact, partnership quality, and bowling pressure metrics will enhance prediction quality.
3. **Deep Learning Models:**
Using LSTM networks, neural networks, or hybrid architectures may capture long-term dependencies in cricket performance.
4. **Real-Time Prediction System:**
A live prediction engine can be built using streaming data during matches to forecast outcomes dynamically.
5. **Player Performance Prediction:**
Future work could focus on predicting individual player runs, wickets, strike rate, or match impact scores.
6. **Interactive Dashboard:**
Building a visual analytics dashboard with tools such as Power BI, Tableau, or Streamlit would allow coaches and analysts to interpret the results more easily.

By addressing these directions, future research can create more comprehensive and actionable analytical tools for ODI cricket performance evaluation.

REFERENCES

- [1] Bunker, R., & Thabtah, F. (2019). *A machine learning framework for sport result prediction: The case of Twenty20 cricket*. *Journal of Sports Analytics*, 5(1), 1–13.
- [2] Kaluarachchi, K. A., & Rohan, T. (2010). *Prediction of ODI cricket matches using machine learning techniques*. *International Journal of Machine Learning and Computing*, 2(4), 273–279.
- [3] Sankaranarayanan, S., Gupta, R., & Sharma, P. (2018). *Deep neural network-based prediction of IPL cricket match results*. *Proceedings of the International Conference on Computational Intelligence*, 221–227.
- [4] Singh, A., Yadav, R., & Patel, K. (2020). *Predicting match outcomes in T20 cricket using ensemble machine learning models*. *Journal of Computational Intelligence in Sports*, 3(2), 55–65.
- [5] Mukherjee, S., & Ravi, V. (2021). *A comparative study of machine learning techniques for cricket player performance modeling*. *Expert Systems with Applications*, 176, 114–124.
- [6] Kumar, S., Sharma, V., & Gupta, A. (2022). *Outcome prediction of T20 cricket matches using logistic regression and random forest*. *International Journal of Data Science Research*, 8(2), 45–53.
- [7] Perera, S., Davis, G., & Swartz, T. (2016). *Modelling T20 cricket outcomes using Bayesian methods*. *Australian & New Zealand Journal of Statistics*, 58(1), 85–101.
- [8] Bailey, M., & Clarke, S. (2017). *Predicting sports scores using Poisson models and machine learning*. *Statistical Modelling in Sports*, 12(3), 122–138.
- [9] Akhtar, M., Scarf, P., & Ruiz, F. (2015). *Forecasting match results in One Day International cricket using GARCH models*. *Journal of Quantitative Analysis in Sports*, 11(1), 35–52.
- [10] Lemmer, H. (2011). *The effectiveness of batting and bowling measures in T20 cricket*. *South African Journal for Research in Sport*, 33(2), 25–36.
- [11] Bhattacharjee, D., & Saikia, H. (2014). *Performance evaluation of cricket players using decision trees*. *International Journal of Sports Science and Engineering*, 8(3), 142–152.


- [12] Shah, A., & Zafar, M. (2019). *Machine learning-based analysis of player performance in T20 cricket*. International Journal of Computer Applications, 178(45), 12–19.
- [13] Keith, L. & Taylor, J. (2018). *T20 cricket match prediction using gradient boosting techniques*. Journal of Applied Data Science, 2(4), 225–240.
- [14] Swartz, T., & Gill, P. (2019). *Cricket analytics: exploring factors that influence match outcomes in T20*. Journal of Sports Analytics, 6(2), 87–101.
- [15] Bhandari, R., & Singh, P. (2020). *A feature engineering approach for predicting cricket match results using machine learning*. International Journal of Pattern Recognition and AI, 34(7), 2056011.
- [16] Ahmed, S., & Chowdhury, I. (2021). *Predictive modelling for Bangladesh cricket using machine learning classification*. Journal of Data Analytics and Sports Informatics, 4(1), 1–12.
- [17] Raj, S., & Kumar, N. (2019). *Analysis of bowling performance in T20 cricket using decision trees*. International Journal of Advanced Computing, 10(5), 45–53.
- [18] Tachafi, Y., & Verma, D. (2022). *An empirical evaluation of ensemble learning techniques for predicting cricket match outcomes*. Annals of Computer Science and Information Systems, 28, 311–318.
- [19] George, A., & Srinivasan, R. (2023). *Applying XGBoost and random forests to predict T20 cricket performance*. Computational Intelligence in Sports Technologies, 5(1), 23–39.
- [20] Jayawardene, K., & Silva, P. (2020). *Impact of toss, venue, and innings on T20 cricket outcomes: A machine learning perspective*. Journal of Sports Data Mining, 8(3), 110–119.
- [21] Thilakarathna, G. N. G. D. H., Navanathara, R. S., Herath, H. M. P. M., Senavirathne, H. T. S. B., Abeysekara, K. D., Mihisara, W. A. S. S. H., Nimsara, W. P. R., Karunarathne, U. S., Wickramasinghe, R. D. R. P., Samarasinghe, W. H. H. H., & Dilshan, L. D. (2025). Statistical analysis of performance of Sri Lankan men's cricket team since 2015. Department of Statistics and Computer Science, University of Kelaniya

- [22]ajib, A. H., Limon, S. H., Naser, A. I., & Saha, G. (2023). Analysing factors impacting Bangladesh men's T20 cricket performance. *Scientific Journal of Sport and Performance*, *3*(1), 79-94.<https://doi.org/10.55860/ENNL4841>
- [23]hilakarathne, J. K. P. M. (2024). Predicting players' performance in 120 cricket matches using machine learning approaches [Undergraduate thesis, University of Jaffna].
- [24]Mahbub, M. K., Miah, M. A. M., Islam, S. M. S., Sorna, S., Hossain, S., & Biswas, M. (2021). Best eleven forecast for Bangladesh cricket team with machine learning techniques. 2021 IEEE. R<https://doi.org/10.1109/SCETECTS0965.2021.3657050>
- [25]Mullick, P. (2024). Statistical evolution of ODI cricket: Analyzing performance trends and effect sizes [Unpublished manuscript]. Department of Operations Research and Business Intelligence, Wroclaw University of Science and Technology.
- [26]Connor, J. D., Sinclair, W. H., Leicht, A. S., & Doma, K. (2019). Analysis of cricket ball type and innings on state level cricket batter's performance. *Frontiers in Psychology*, 10, 2347.<https://doi.org/10.3389/fpsyg.2019.02347>
- [27]Mullick, P. (2024). Statistical evolution of ODI cricket: Analyzing performance trends and effect sizes. Department of Operations Research and Business Intelligence, Wroclaw University of Science and Technology, Poland.
- [28]Mukherjee, S. (2014). Quantifying individual performance in cricket-A network analysis of batsmen and bowlers. *Physica A: Statistical Mechanics and its Applications*, 393, 624-637.<https://doi.org/10.1016/j.physa.2013.09.027>
- [29]Shah, S. A. A., Zaman, Q., Qazi, U.-U.-S., Iqbal, T., & Mubarak, A. (2024). Influential factors in successful run chases in T20 International cricket: A statistical and predictive analysis. *International Journal of Politics & Social Sciences Review*, 3(III), 20-29
- [30]Thilakarathne, J. K. P. M. (2024). Predicting players' performance in T20 cricket matches using Machine Learning approaches [Thesis]. University of Jaffna.
- [31]Sadekar, O., Chowdhary, S., Santhanam, M. S., & Battiston, F. (2024). Individual and team performance in cricket. *Royal Society Open Science*, 11, 240809.<https://doi.org/10.1098/rsos.240809>

Md. Amanullah Aman

221-35-820

 Quick Submit

 Quick Submit

 Daffodil International University

Document Details

Submission ID

trn:oid::1:3450438867

Submission Date

Dec 24, 2025, 11:50 AM GMT+6

Download Date

Dec 24, 2025, 12:53 PM GMT+6

File Name

BD_Cricket_performance_ID_221-35-820.pdf

File Size

549.3 KB

43 Pages

7,374 Words

44,462 Characters

*% detected as AI

AI detection includes the possibility of false positives. Although some text in this submission is likely AI generated, scores below the 20% threshold are not surfaced because they have a higher likelihood of false positives.

Caution: Review required.

It is essential to understand the limitations of AI detection before making decisions about a student's work. We encourage you to learn more about Turnitin's AI detection capabilities before using the tool.

Disclaimer

Our AI writing assessment is designed to help educators identify text that might be prepared by a generative AI tool. Our AI writing assessment may not always be accurate (i.e., our AI models may produce either false positive results or false negative results), so it should not be used as the sole basis for adverse actions against a student. It takes further scrutiny and human judgment in conjunction with an organization's application of its specific academic policies to determine whether any academic misconduct has occurred.

Frequently Asked Questions

How should I interpret Turnitin's AI writing percentage and false positives?

The percentage shown in the AI writing report is the amount of qualifying text within the submission that Turnitin's AI writing detection model determines was either likely AI-generated text from a large-language model or likely AI-generated text that was likely revised using an AI paraphrase tool or word spinner.

False positives (incorrectly flagging human-written text as AI-generated) are a possibility in AI models.

AI detection scores under 20%, which we do not surface in new reports, have a higher likelihood of false positives. To reduce the likelihood of misinterpretation, no score or highlights are attributed and are indicated with an asterisk in the report (*%).

The AI writing percentage should not be the sole basis to determine whether misconduct has occurred. The reviewer/instructor should use the percentage as a means to start a formative conversation with their student and/or use it to examine the submitted assignment in accordance with their school's policies.



What does 'qualifying text' mean?

Our model only processes qualifying text in the form of long-form writing. Long-form writing means individual sentences contained in paragraphs that make up a longer piece of written work, such as an essay, a dissertation, or an article, etc. Qualifying text that has been determined to be likely AI-generated will be highlighted in cyan in the submission, and likely AI-generated and then likely AI-paraphrased will be highlighted purple.

Non-qualifying text, such as bullet points, annotated bibliographies, etc., will not be processed and can create disparity between the submission highlights and the percentage shown.

13%

SIMILARITY INDEX

11%

INTERNET SOURCES

4%

PUBLICATIONS

7%

STUDENT PAPERS

PRIMARY SOURCES

1	Submitted to Midlands State University Student Paper	4%
2	Submitted to NCC Education Student Paper	1%
3	Pushpa Choudhary, Sambit Satpathy, Arvind Dagur, Dharendra Kumar Shukla. "Recent Trends in Intelligent Computing and Communication", CRC Press, 2025 Publication	1%
4	indah.ump.edu.my Internet Source	1%
5	Sukhpreet Kaur, Amanpreet Kaur, Manish Kumar. "Recent Advances in Computational Methods in Science and Technology - Volume 2", CRC Press, 2026 Publication	<1%
6	umpir.ump.edu.my Internet Source	<1%
7	Submitted to National Taiwan University Student Paper	<1%
8	Submitted to Daffodil International University Student Paper	<1%
9	Submitted to Visvesvaraya Technological University Student Paper	<1%
10	www.researchgate.net Internet Source	<1%

11	pmc.ncbi.nlm.nih.gov Internet Source	<1 %
12	R. N. V. Jagan Mohan, B. H. V. S. Rama Krishnam Raju, V. Chandra Sekhar, T. V. K. P. Prasad. "Algorithms in Advanced Artificial Intelligence - Proceedings of International Conference on Algorithms in Advanced Artificial Intelligence (ICAAAI-2024)", CRC Press, 2025 Publication	<1 %
13	www.coursehero.com Internet Source	<1 %
14	Submitted to Dublin Business School Student Paper	<1 %
15	www.imse.iastate.edu Internet Source	<1 %
16	www.ijecs.in Internet Source	<1 %
17	www.viirj.org Internet Source	<1 %
18	Ebenezer Fiifi Emire Atta Mills, Zihui Deng, Zhuoqing Zhong, Jinger Li. "Data-driven prediction of soccer outcomes using enhanced machine and deep learning techniques", Journal of Big Data, 2024 Publication	<1 %
19	docshare.tips Internet Source	<1 %
20	wseas.com Internet Source	<1 %
21	www.frontiersin.org Internet Source	<1 %

22	Submitted to Indian Institute of Technology Guwahati Student Paper	<1 %
23	dspace.bracu.ac.bd Internet Source	<1 %
24	ijirt.org Internet Source	<1 %
25	ijsrst.com Internet Source	<1 %
26	repository.unsri.ac.id Internet Source	<1 %
27	researchmgt.monash.edu Internet Source	<1 %
28	rosap.ntl.bts.gov Internet Source	<1 %
29	storage.googleapis.com Internet Source	<1 %
30	Iqbal, Mazhar. "Physical Sensing for Assessing Environmental Influences on Bird Distribution, Migration, and Demography: An Integrated Study of Avian Diversity, Vocalization Patterns, and Habitat Characteristics", The University of Texas at Dallas Publication	<1 %
31	dspace.daffodilvarsity.edu.bd:8080 Internet Source	<1 %
32	Al Falasi, Bakhit. "Dubai Urban Planning 2040: Resource Requirements in the Urban District 5.", Rochester Institute of Technology, 2024 Publication	<1 %
33	T. Joshva Devadas, Naveenkumar Jayakumar, N. Narayanan Prasanth, P. Anandhakumar,	<1 %

Shrikant Jadhav. "Advanced Computational Intelligence Techniques for Engineers", CRC Press, 2025

Publication

34	download.bibis.ir Internet Source	<1 %
35	export.arxiv.org Internet Source	<1 %
36	matthewstephenson.info Internet Source	<1 %
37	mobt3ath.com Internet Source	<1 %
38	text2fa.ir Internet Source	<1 %
39	thesis.eur.nl Internet Source	<1 %
40	uis.brage.unit.no Internet Source	<1 %

Exclude quotes Off

Exclude matches Off

Exclude bibliography Off

Account clearance

The screenshot displays a student dashboard with a dark sidebar on the left containing navigation options: Dashboard, Student Profile, Payment Ledger, Registrations/Exam Clearance, Registered Course, Result, Routing, Live Result, Teaching Evaluation, Scholarship, and Convocation Apply. The main content area is titled 'Dashboard' and 'Student Portal'. It features four blue summary cards: 'Total Payable' (767,200.00), 'Total Paid' (767,200.00), 'Total Due' (0.00), and 'Total Other' (2,925.00). Below these is a section for 'Today's Routine - Thursday' with a note that no routine is available for today. At the bottom, there is a section for 'Semester Wise Result' with a link to 'Semester-wise SGPA Performance'. The user's name 'MO. AMANULLAH AMAN' and ID '221-55-830' are visible in the top right corner.

Category	Amount
Total Payable	767,200.00
Total Paid	767,200.00
Total Due	0.00
Total Other	2,925.00