

**A
Project Report
on
Multi-model Approach for IPL Match Prediction and Team
Selection**

**Submitted to
Sant Gadge Baba Amravati University, Amravati**

**Submitted in partial fulfilment of
the requirements for the Degree of
Bachelor of Engineering in
Computer Science and Engineering**

Submitted by

Gaurav B. Wankhede

(PRN:213120225)

Krishna M. Kashikar

(PRN:213120350)

Aniket V. Rajguru

(PRN:213120154)

Prashik T. Ingle

(PRN:213120258)

Jivan S. Thorat

(PRN: 203120331)

**Under the Guidance of
Prof. Ms. Pooja R. Pohare
Asst.Prof. CSE Dept.**



**Department of Computer Science and Engineering
Shri Sant Gajanan Maharaj College of Engineering,
Shegaon – 444 203 (M.S.)**

Session 2024-2025

**SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,
SHEGAON – 444 203 (M.S.)**

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



CERTIFICATE

This is to certify that **Mr. Gaurav B. Wankhede, Mr. Aniket V. Rajguru, Mr. Krishna M. Kashikar, Mr. Prashik T. Ingle and Mr. Jivan S. Thorat** students of final year Bachelor of Engineering in the academic year 2024-25 of Computer Science and Engineering Department of this institute have completed the project work entitled “**Multi-Model Approach For IPL Match Prediction And Team Selection**” and submitted a satisfactory work in this report. Hence recommended for the partial fulfilment of degree of Bachelor of Engineering in Computer Science and Engineering.

Prof. Ms. Pooja R. Pohare
Project Guide

Dr. J. M. Patil
Head of Department

Dr. S. B. Somani
Principal
SSGMCE, Shegaon


**SHRI SANT GAJANAN MAHARAJ COLLEGE OF ENGINEERING,
SHEGAON – 444 203 (M.S.)**


DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING



CERTIFICATE

This is to certify that **Mr. Gaurav B. Wankhede, Mr. Aniket V. Rajguru, Mr. Krishna M. Kashikar, Mr. Prashik T. Ingle and Mr. Jivan S. Thorat** students of final year Bachelor of Engineering in the academic year 2024-25 of Computer Science and Engineering Department of this institute have completed the project work entitled “**Multi-Model Approach For IPL Match Prediction And Team Selection**” and submitted a satisfactory work in this report. Hence recommended for the partial fulfillment of degree of Bachelor of Engineering in Computer Science and Engineering.


Internal Examiner
Prof. P. R. Pohare
Name and Signature
Date: 9/5/25


External Examiner
A R Ladole.
Name and Signature
Date:

Acknowledgement

It is our utmost duty and desire to express gratitude to various people who have rendered valuable guidance during our project work. We would have never succeeded in completing our task without the cooperation, encouragement and help provided to us by them. There are several people who deserve recognition for their unwavering support and guidance throughout this report.

We are highly indebted to our guide **Prof. Ms. Pooja R. Pohare** for his guidance and constant supervision as well as for providing necessary information from time to time. We would like to take this opportunity to express our sincere thanks, for his esteemed guidance and encouragement. His suggestions broaden our vision and guided us to succeed in this work.

We are sincerely thankful to **Dr. J. M. Patil** (HOD, CSE Department, SSGMCE, Shegaon), and to **Dr. S. B. Somani** (Principal, SSGMCE, Shegaon) who always has been kind to extend their support and help whenever needed.

We would like to thank all teaching and non-teaching staff of the department for their cooperation and help. Our deepest thank to our parents and friends who have consistently assisted us towards successful completion of our work.

-Projectors

Gaurav B. Wankhade (22BD310748)

Aniket V. Rajguru (22BD310735)

Krishna M. Kashikar (22BD310758)

Prashik T. Ingle (22BD310774)

Jivan S. Thorat (21BD310742)

Abstract	i
List of Figures	ii
List of Screenshots	iii
1. Introduction	1
1.1 Preface	1
1.2 Motivation	2
1.3 Problem Statements	3
1.4 Objectives	4
1.5 Scope of Project	5
1.6 Organization of Project	6
2. Literature Review	7
2.1 Integer Optimization for Dream 11 Cricket Team Selection	7
2.2 ICC T20 Cricket World Cup 2020 Winner Prediction Using Machine Learning Techniques	7
2.3 Best Eleven Forecast for Bangladesh Cricket Team with Machine Learning Techniques	8
2.4 Cricket Team Prediction Using Machine Learning Techniques	8
2.5 Literature Survey Summery	9

3. Methodology	11
3.1 Methodology	11
3.2 Proposed System	13
3.3 System Model	14
3.4 Algorithms Used	16
4. Implementation	20
5. Result and Discussion	25
6. Conclusion	32
6.1 Conclusion	32
6.2 Future Scope	32
References	34
Dissemination of Work	36
Plagiarism Report	37
Project Group Members	42

Abstract

This project presents a multi-model machine learning approach for predicting the best XI players for fantasy cricket on the Dream11 platform. The system leverages player performance metrics such as runs and wickets to develop accurate prediction models using Random Forest and Decision Tree algorithms. The objective is to enhance the user experience by offering data-driven recommendations for team selection, increasing the chances of success in fantasy leagues. The project includes secure user authentication, dataset preprocessing, model training, evaluation, and visualization. With a focus on automation and improved prediction accuracy, this solution significantly reduces the time required for manual analysis. It demonstrates the practical application of machine learning in sports analytics, highlighting the potential for further development in the domain of fantasy sports.

Keywords: Dream11, Fantasy Cricket, Machine Learning, Random Forest, Decision Tree, Player Prediction, Sports Analytics, Team Selection, Python, Data Preprocessing

List of Figures

Figure No.	Description
Figure 3.1	Proposed System Model
Figure 3.2	Working Flow of Model
Figure 3.3	Working of Random Forest Algorithm
Figure 3.4	Working of Logistic Regression Classifier
Figure 3.5	Working of Linear Regression Algorithm
Figure 4.1	Real Implementation Flow Diagram
Figure 4.2	Simple app.py File

List of Screenshots

Screenshot No.	Description
Screenshot 5.1	Authentication Page
Screenshot 5.2	Home Screen Page
Screenshot 5.3	First Inning Score Prediction1
Screenshot 5.4	First Inning Score Prediction2
Screenshot 5.5	First Inning Score Prediction3
Screenshot 5.6	IPL Match Win Prediction1
Screenshot 5.7	IPL Match Win Prediction2
Screenshot 5.8	IPL Match Win Prediction3
Screenshot 5.9	IPL Team Prediction1
Screenshot 5.10	IPL Team Prediction2
Screenshot 5.11	IPL Team Prediction3
Screenshot 5.12	Model Accuracy

CHAPTER 1
INTRODUCTION

1. INTRODUCTION

1.1 PREFACE

Fantasy sports have gained immense popularity in recent years, particularly in cricket-loving countries like India. Among the leading platforms in this space is Dream11, which allows users to create virtual teams comprising real-life players. Users earn points based on the actual performances of the selected players during live matches. With millions of users participating daily, making accurate predictions and smart player selections has become essential for success [4].

This project aims to develop an intelligent and automated system that predicts the best eleven players for any given match using machine learning (ML) algorithms. By analyzing historical performance data—such as runs scored, wickets taken, and consistency metrics—this system helps users make informed decisions while selecting their teams [1]. The implementation of Random Forest and Decision Tree classifiers ensures that player predictions are not only statistically robust but also tailored to the dynamics of each match.

The primary motivation behind this research is to enhance the user experience on fantasy platforms by reducing the time spent on manual analysis and increasing the accuracy of team selection [7]. Additionally, the proposed model can be extended to other fantasy sports formats, making it a versatile solution in the domain of sports analytics.

This project integrates various components such as data collection, preprocessing, feature selection, and model training using Python libraries like NumPy, Pandas, and Scikit-learn. The final model evaluates different combinations of players and recommends the most optimal fantasy XI for a given match scenario.

By combining domain knowledge from cricket with the power of machine learning, this project not only provides a practical application but also sets a foundation for future advancements in AI-based fantasy sports management.

1.2 MOTIVATION

Cricket is not just a sport but a passion for millions, especially in India. With the rise of digital platforms, Dream11 has revolutionized the way fans engage with the game by allowing them to create fantasy teams and compete based on real match performances. As the competition intensifies with millions of users, the need for data-driven decision-making has become crucial.

Currently, most users rely on personal judgment or basic statistics to form their teams, which often leads to inconsistent results [5]. This project is motivated by the idea of using machine learning techniques to assist users in selecting the most optimal team with higher chances of winning. By leveraging historical performance data and predictive algorithms like Random Forest and Decision Tree, users can make smarter, faster, and more accurate selections [2].

Additionally, the project aims to reduce the time and effort required in researching player statistics by providing an automated and intelligent team recommendation system. This not only improves user satisfaction but also adds a layer of strategic depth to the fantasy gaming experience.

1.3 PROBLEM STATEMENT

In the rapidly growing domain of fantasy sports, especially on platforms like Dream11, users are often faced with the challenge of selecting the best eleven players from two competing teams. This selection is typically based on manual analysis or limited statistical data, which may lead to suboptimal decisions and lower success rates.

There is a need for an intelligent system that can analyze historical player performance data and predict the most suitable players for team selection. The absence of a reliable, automated solution limits the user's ability to make informed choices, affecting both user engagement and competition outcomes.

This project proposes a machine learning-based model that predicts the best XI players using algorithms like Random Forest and Decision Tree, with the goal of improving team selection accuracy, enhancing user experience, and increasing win probabilities in fantasy cricket leagues.

1.4 OBJECTIVES

1. Develop a Predictive Model:

Build a machine learning-based model to predict the best XI players for fantasy cricket based on historical performance data.

2. Enhance Team Selection Accuracy:

Improve the accuracy of fantasy team selection by using reliable algorithms like Random Forest and Decision Tree.

3. Simplify the User Experience:

Provide users with an easy-to-use system that offers intelligent recommendations, reducing the need for manual research.

4. Implement Secure User Access:

Develop a secure and user-friendly login and registration system to manage access to the platform.

5. Model Evaluation and Optimization:

Compare different algorithms, analyze performance, and optimize the most accurate model for player prediction.

6. Publish Research Outcomes:

Share the findings through publication in reputed journals or conferences to contribute to the academic and professional community.

1.5 SCOPE OF PROJECT

Fantasy Cricket Team Selection:

The system is specifically designed for fantasy cricket team selection, focusing on Dream11 matches and player performance.

Use of Machine Learning Algorithms:

The project employs machine learning models such as Random Forest and Decision Tree for accurate predictions.

Historical Data Analysis:

The model utilizes historical player statistics (e.g., runs, wickets, averages) for generating predictions.

Automated Player Recommendation:

Offers automated suggestions for the best eleven players, saving users time and effort.

Scalability:

Though developed for IPL and Dream11, the system can be extended to other cricket formats or fantasy sports with minimal changes.

User Interface:

Includes a basic, secure interface for user authentication and accessing prediction features.

1.6 ORGANIZATION OF PROJECT

The project is systematically structured into six chapters, each focusing on a particular aspect of the research and development process. The purpose of this organization is to provide a clear and logical flow from the initial concept to the final outcome, ensuring that the reader gains a comprehensive understanding of the entire project lifecycle.

Chapter 1: Introduction

This chapter introduces the background and context of the project. It covers the rise of fantasy sports platforms like Dream11 and highlights the need for a data-driven approach to player selection. The chapter outlines the motivation, problem statement, objectives, and scope of the project. It also discusses the significance of integrating machine learning into fantasy sports for improving user engagement and team selection accuracy.

Chapter 2: Literature Survey

The literature survey explores previous research related to player prediction in fantasy sports and the application of machine learning in cricket analytics. Various techniques and models used in earlier studies—such as Random Forest, Naive Bayes, and Knapsack algorithms—are reviewed and compared. This chapter helps establish a theoretical foundation and justifies the use of specific algorithms selected for this project.

Chapter 3: System Analysis and Design

This chapter presents a detailed analysis of the system's architecture and design. It outlines the functional and non-functional requirements, data flow diagrams, system workflow, and use case diagrams. The chapter also explains how the system handles data collection, preprocessing, model training, and team prediction. The methodology followed in this chapter ensures the logical and technical feasibility of the project.

Chapter 4: Implementation

The implementation chapter details the practical execution of the project. It includes the tools and technologies used, such as Python, NumPy, Pandas, and Scikit-learn. The chapter explains the steps for dataset preparation, feature selection, model training, and model testing. The integration of machine learning models like Random Forest and

Decision Tree is explained with code-level insights and workflow diagrams. This chapter transforms the design concepts into a working system.

Chapter 5: Results and Discussion

In this chapter, the experimental results of the implemented models are presented. It includes performance metrics such as accuracy, precision, recall, and F1-score. Comparisons between the algorithms are made to determine which model performs best in terms of predictive accuracy for selecting the Dream11 fantasy XI. Visualizations and graphs are used to help interpret the outcomes effectively. This chapter also discusses any limitations and unexpected findings during model evaluation.

Chapter 6: Conclusion and Future Work

The final chapter summarizes the key findings and contributions of the project. It reflects on how the proposed system meets the defined objectives and discusses the impact of using machine learning in fantasy cricket applications. The chapter concludes with suggestions for future enhancements, such as integrating deep learning models, expanding the system to other sports, and deploying the system as a live web application or mobile app.

CHAPTER 2
LITERATURE
REVIEW

2. LITERATURE REVIEW

2.1 Integer Optimization for Dream 11 Cricket Team Selection

Authors: Saurav Singla, Swapna Samir Shukla, International Journal of Computer Sciences and Engineering

This paper explores the use of integer optimization techniques for selecting the ideal Dream11 cricket team. The authors focus on optimizing player selection based on various performance metrics while considering constraints such as team budget, player roles (batsman, bowler, all-rounder), and the maximum number of players from a single team. The approach leverages optimization algorithms to propose an efficient team selection strategy that maximizes the overall points for a fantasy cricket game [1]

2.2 ICC T20 Cricket World Cup 2020 Winner Prediction Using Machine Learning Techniques

Authors: Abdul Basit, Rehan Ali Shah, *IEEE*

This paper investigates the use of machine learning techniques to predict the winner of the ICC T20 World Cup 2020. The authors apply various classification algorithms, such as Random Forest, Support Vector Machines, and Logistic Regression, to build predictive models based on historical match data, player performance, and team dynamics. The goal is to determine the factors influencing a team's chances of winning and use this data to accurately predict outcomes in the tournament [2].

2.3 Best Eleven Forecast for Bangladesh Cricket Team with Machine Learning Techniques

Authors: Md. KawsherMahbub, Md. Abdul MozidMiah, IEEE

This paper aims to predict the best playing eleven for the Bangladesh cricket team using machine learning techniques. The authors analyze historical performance data, player statistics, fitness levels, and other relevant features to create a model that suggests the most optimal team composition for any given match. The approach combines data preprocessing, feature extraction, and machine learning algorithms like Random Forest and Logistic Regression to forecast the best eleven players based on match conditions and team strategies [3].

2.4 Cricket Team Prediction Using Machine Learning Techniques

Authors: Nilesh M. Patil, Bevan H. Sequeira, SSRN 2021

In this paper, the authors explore the use of machine learning to predict cricket team performance and match outcomes. They use various classification and regression techniques to analyze historical player data, team performance, and match conditions. The study focuses on building models that can predict the winning probability of a team, based on factors like player form, team strength, weather conditions, and other match variables. The paper highlights the application of data-driven decision-making for team management and match predictions [4].

2.5 LITERATURE SURVEY SUMMARY

Dream11 Cricket Team Selection:

Uses integer optimization to maximize fantasy points while adhering to budget and player role constraints.

ICC T20 World Cup 2020 Prediction:

Applies machine learning models like Random Forest and SVM to predict match outcomes based on historical data.

Best Eleven for Bangladesh Cricket:

Predicts the optimal playing eleven using machine learning by analyzing player performance and fitness.

Cricket Team Prediction:

Uses machine learning to predict team performance and match outcomes based on historical and external data.

CHAPTER 3
METHODOLOGY

3 METHODOLOGY

3.1 Methodology used in this Project

3.1.1 Data Collection and Preprocessing:

Data Acquisition: Gather historical data for players, teams, and matches, including player performance stats (runs, wickets, fitness, etc.), team performance, match conditions, and external factors (weather, pitch report, etc.).

Data Cleaning: Handle missing data, outliers, and irrelevant information by applying appropriate techniques such as imputation, normalization, and standardization.

Feature Engineering: Extract important features such as player form, team dynamics, and match-specific conditions (e.g., pitch type, weather) to enhance the model's predictive power.

3.1.2 Data Integration:

Multimodal Data Combination: Combine structured data (player stats, historical performance) with unstructured data (match reports, player fitness) to form a comprehensive dataset for analysis.

Data Transformation: Transform raw data into usable formats suitable for machine learning algorithms, including encoding categorical variables and scaling numerical features.

3.1.3 Model Selection and Training:

Algorithm Selection: Implement multiple machine learning algorithms such as:

Random Forest: For classification and regression tasks, especially in predicting match outcomes.

Logistic Regression: To predict winning probabilities based on player and team data.

Naive Bayes: For predicting match outcomes based on historical data.

Decision Trees: For interpreting feature importance and team selection strategies.

Linear Regression: To predict continuous variables such as player performance or match scores.

Model Training: Split the dataset into training and testing sets, and train the models using cross-validation techniques to prevent over fitting and ensure robustness.

3.1.4 Prediction and Evaluation:

Prediction: Use the trained models to predict:

The best playing eleven for cricket teams based on player statistics, fitness, and match conditions.

The match winner by evaluating team performance, player form, and other factors.

Evaluation Metrics: Evaluate the performance of the models using metrics such as accuracy, precision, recall, F1-score, and confusion matrices.

Model Comparison: Compare the performance of different models and choose the one with the highest predictive accuracy for team selection and match prediction.

3.1.5 Results and Insights:

Analysis: Analyze the results from the prediction models to identify key factors influencing match outcomes and team performance.

Insights: Provide actionable insights for improving player selection and team strategies, offering valuable decision-making tools for coaches, analysts, and fantasy sports platforms.

3.1.6 Deployment (Optional):

Develop a user-friendly application or web platform that allows users (e.g., fantasy sports players, coaches) to input data and get real-time predictions for team selection and match outcomes.

This methodology combines data preprocessing, machine learning modeling, optimization, and evaluation to achieve effective player selection and match outcome prediction, ensuring robust and accurate results for sports analysis.

3.2 Proposed System

The proposed system is a machine learning-based fantasy cricket team predictor that assists users in selecting the best eleven players for a Dream11 match. The system analyzes historical performance data of players and uses predictive modeling techniques to recommend an optimal team based on recent form, consistency, and overall performance metrics. The key idea is to automate the team selection process using well-known supervised learning algorithms such as Random Forest and Decision Tree, which have been proven effective in sports prediction tasks. These models are trained on a dataset that includes features such as runs scored, wickets taken, matches played, batting and bowling averages, strike rates, and economy rates [2].

3.3 System Model

Here's a **workflow** for the proposed system:

1. **Data Collection:** Gather data on players, teams, matches, and external factors.
2. **Data Preprocessing:** Clean, transform, and engineer features from raw data.
3. **Model Selection:** Choose machine learning models (Random Forest, Logistic Regression, Naive Bayes, etc.).
4. **Model Training:** Train models using historical data and validate performance.
5. **Model Optimization:** Fine-tune hyper parameters using Grid Search.
6. **Prediction:** Predict the best playing eleven and match outcomes.
7. **Evaluation:** Evaluate models' performance using metrics like accuracy and F1-score.
8. **Output:** Provide player selection recommendations and match predictions.
9. **Deployment (Optional):** Develop an application for real-time predictions.

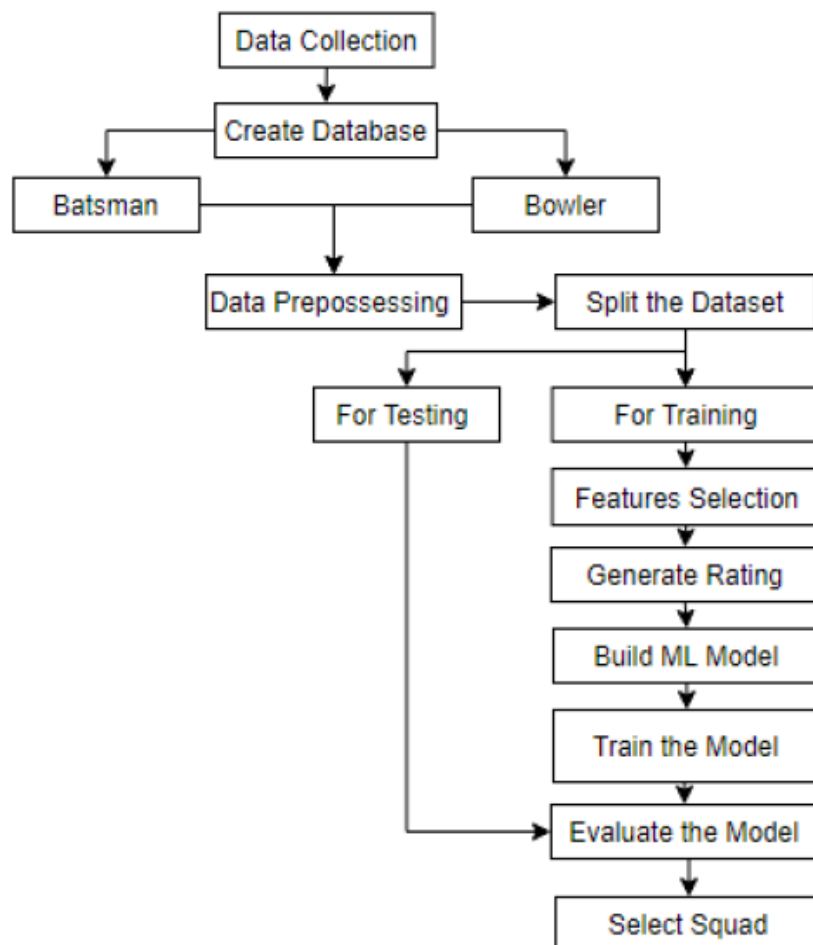


Figure 3.1 Proposed System Model

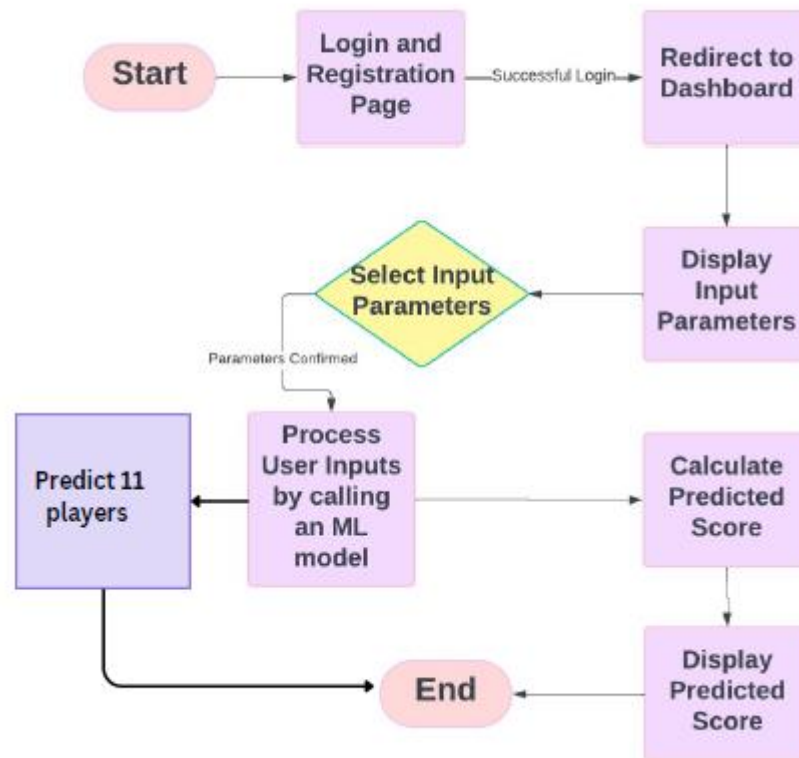


Figure 3.2 Working Flow of Model

3.4 ALGORITHMS USED IN THIS PROJECT

Here's a brief overview of the algorithms applied in the project for IPL player selection and match prediction, along with their roles:

1. Random Forest Algorithm:

- **Type:** Ensemble Learning (Classification and Regression)
- **Purpose:** To predict match outcomes and select the best players for a team.
- **How It Works:**

Random Forest constructs a collection of decision trees using random subsets of the data and features. It then aggregates the predictions from all trees (majority voting for classification or averaging for regression) to make the final decision.

It reduces overfitting by averaging multiple decision trees and handles both classification (win/loss prediction) and regression (predicting player performance) tasks.

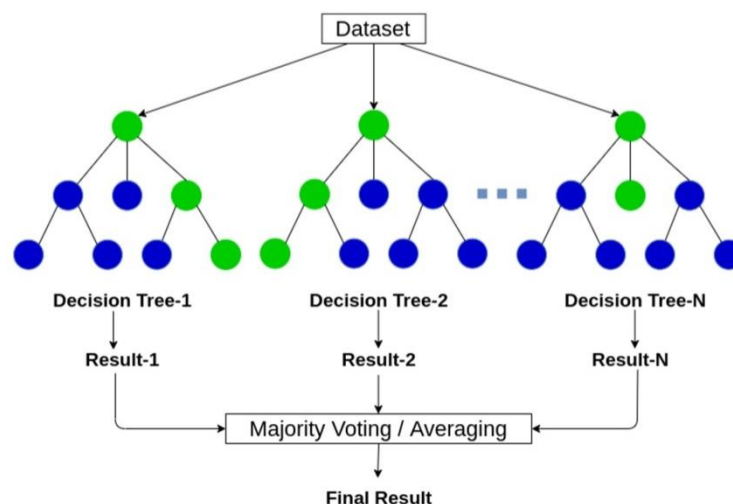


Figure 3.3 Working of Random Forest Algorithm

2. Logistic Regression:

- **Type:** Supervised Learning (Classification)
- **Purpose:** To predict the probability of a binary outcome, such as match results (win/loss).
- **How It Works:**

Logistic regression models the probability that a given input (player stats, team performance, etc.) belongs to a particular class (e.g., win or loss).

It uses the logistic function (sigmoid curve) to map predictions between 0 and 1, making it suitable for binary classification.

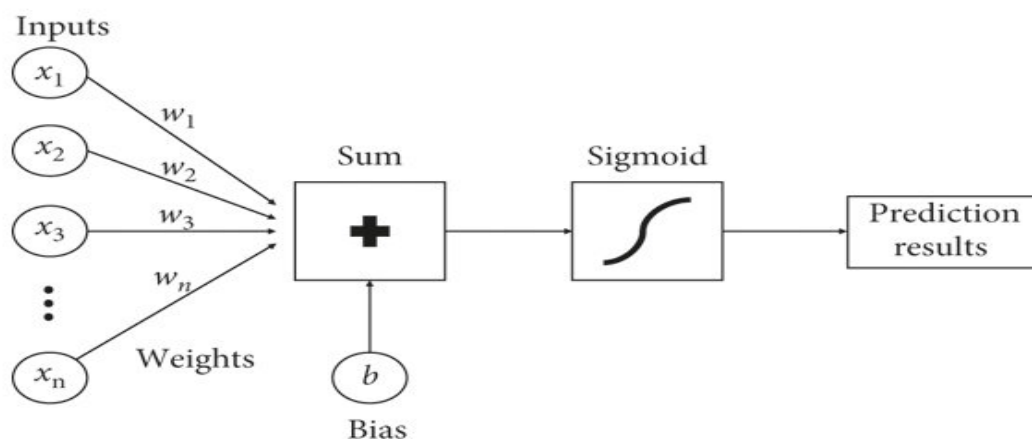


Figure 3.4 Working of Logistic Regression Classifier

3. Linear Regression:

- **Type:** Supervised Learning (Regression)
- **Purpose:** To predict continuous values such as player scores or match points.
- **How It Works:**

Linear regression models the relationship between independent variables (e.g., player stats) and a dependent variable (e.g., match score) by fitting a linear equation to observed data. It minimizes the difference between predicted and actual values (i.e., least squares method).

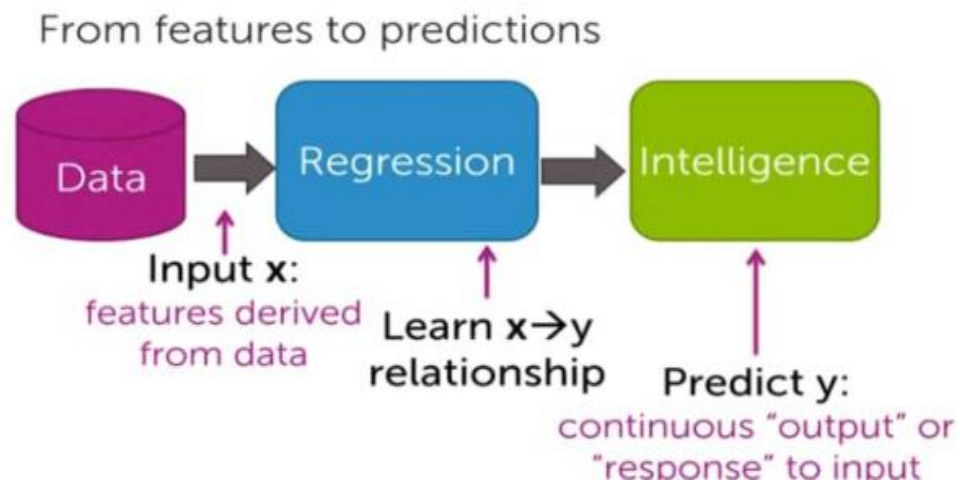


Figure 3.5 Working of Linear Regression Algorithm

Summary of Algorithms Used:

- **Random Forest:** Ensures robust predictions with multiple decision trees, reducing overfitting.
- **Logistic Regression:** Used for binary classification tasks, such as predicting match outcomes.
- **Linear Regression:** Used for continuous predictions like player scores or match points.

These algorithms combined help in accurately predicting match outcomes, optimizing team selection, and improving decision-making processes for sports management and fantasy leagues.

CHAPTER 4
IMPLEMENTATION

4 IMPLEMENTATION

The implementation of this project starts with identifying the core objective: to predict player selection and enhance team composition in the Indian Premier League (IPL), with the potential to adapt the model for other sports like football, hockey, and Pro Kabaddi.

The first step involves gathering a wide range of data, including structured numerical statistics like runs, wickets, strike rates, and player roles. Additional data sources include ball-by-ball match events, match outcomes, environmental conditions such as pitch type and weather, and player-specific factors like fitness level, age, and recent injury history [3]. These datasets can be collected from online platforms such as Kaggle, official IPL websites, sports APIs, and wearable fitness devices.

Once the data is collected, it goes through a thorough preprocessing stage. This includes cleaning missing values, removing duplicates or inconsistencies, and formatting all data uniformly. Feature engineering plays a crucial role at this stage, where new attributes like recent form, weighted averages of performance, and encoded categorical variables (such as team name, venue, and player type) are derived. Normalization and scaling are also applied to ensure all features are on a similar scale, which improves the performance of several machine learning algorithms. Additionally, feature selection methods like correlation analysis and recursive feature elimination help in identifying the most relevant predictors for accurate model building.

The modeling phase uses a combination of machine learning algorithms to predict player selection probabilities. The algorithms include Random Forest, Naïve Bayes, Decision Trees, Logistic Regression, and Linear Regression. Among these, the Random Forest classifier consistently delivers the best results, achieving an accuracy of followed by Naïve Bayes at and Decision Trees The models are trained and validated using train-test splits or cross-validation techniques to ensure they generalize well to unseen data. Hyper parameter tuning using Grid Search is applied to optimize the performance of each model further. After training, the models generate predictions that help rank players based on their

performance potential or likelihood of selection. These rankings are then used to form an optimized team using custom logic that considers role balance (batsmen, bowlers, all-rounders), player form, and other selection constraints like team quotas or salary caps. More advanced optimization techniques like genetic algorithms or greedy algorithms can also be integrated to refine team combinations and adhere to league-specific requirements.

Finally, to make the system user-friendly and interactive, the entire model pipeline can be deployed as a web-based application using frameworks like Flask. This allows users—such as coaches, analysts, or fantasy players—to upload data, receive real-time player predictions, and visualize team compositions. The modular design of the system ensures that with minimal adjustments, the same approach can be adapted for other team-based sports by replacing sport-specific features. This end-to-end implementation provides a data-driven and scalable solution to the complex task of player selection and team optimization in competitive sports.

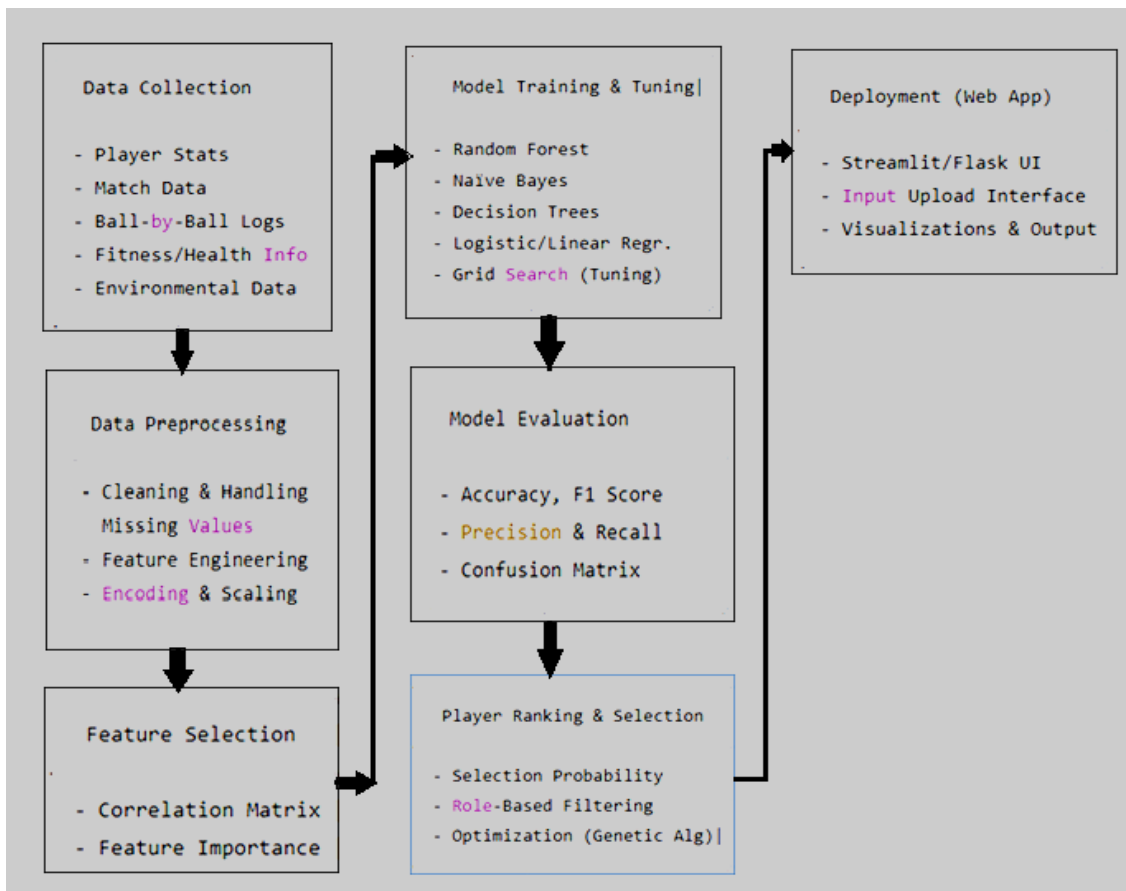


Figure 4.1 Real Implementation Flow Diagram

Implementing IPL Player Selection using Python and Flask involves building a web application where users can select players for two teams and get Dream11 predictions based on historical IPL data using machine learning (e.g., Random Forest). Here's a full breakdown of how it can be done:

Tools and Libraries

- **Backend:** Flask (Python web framework)
- **Frontend:** HTML, CSS, JavaScript
- **ML:** Pandas, Scikit-learn (Random Forest Classifier)
- **Data:** ipl_data.csv and other IPL datasets (historical IPL performance data)

Python and Flask work together to create web applications by allowing Python code to handle requests from web browsers and generate responses, including HTML pages and other web content. Flask provides the framework for handling these requests and responses, while Python serves as the language for writing the logic and code.

Flask is a lightweight and flexible backend web framework for Python. It is used for web development quickly and easily. It is an API of Python that allows us to build web applications. Flask's framework is more explicit than Django's framework and is also easier to learn because it has less base code to implement a simple web application. Flask Python is based on the WSGI (Web Server Gateway Interface) toolkit and Jinja2 template engine.

```
from flask import Flask
app = Flask(__name__) # Flask constructor

# A decorator used to tell the application
# which URL is associated function
@app.route('/')
def hello():
    return 'HELLO'

if __name__ == '__main__':
    app.run(debug=True)
```

Figure 4.2 Simple app.py file

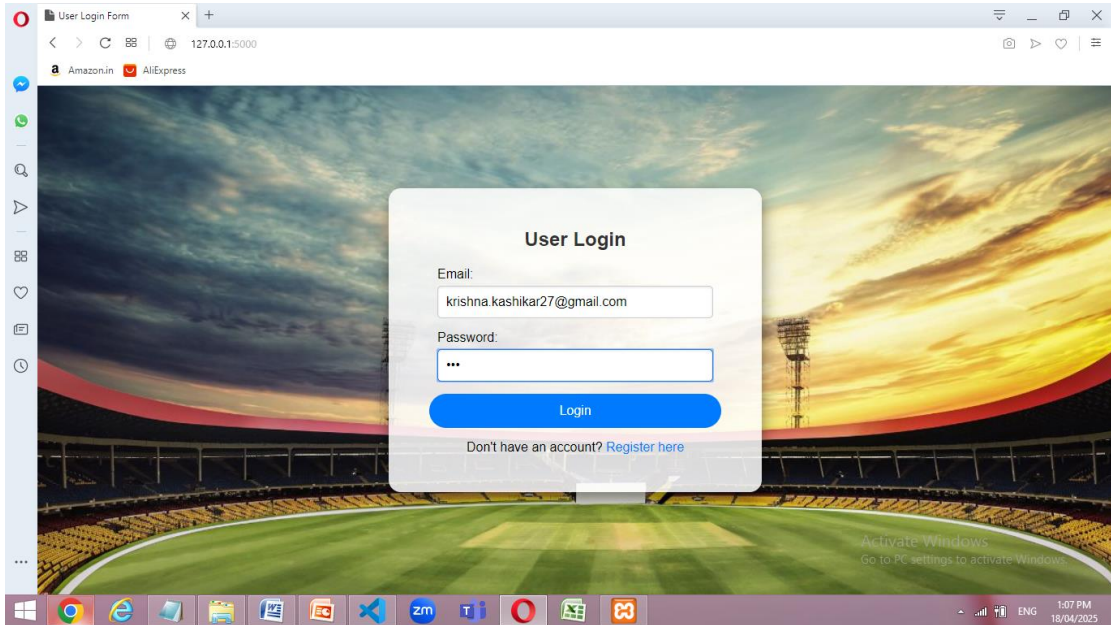
Explanation:

- **Flask(__name__)**: Creates the Flask app.
- **@app.route('/')**: Defines the home route (/).
- **def hello()**: creates a function that is bound with '/' route and returns "HELLO" when the root page is accessed.
- **app.run(debug=True)**: runs the app in debug mode. It ensures that app is not need to restart manually if any changes are made in code.
- Use python app.py to run this program
- From the browser run <http://127.0.0.1:5000> in order to open the web app.

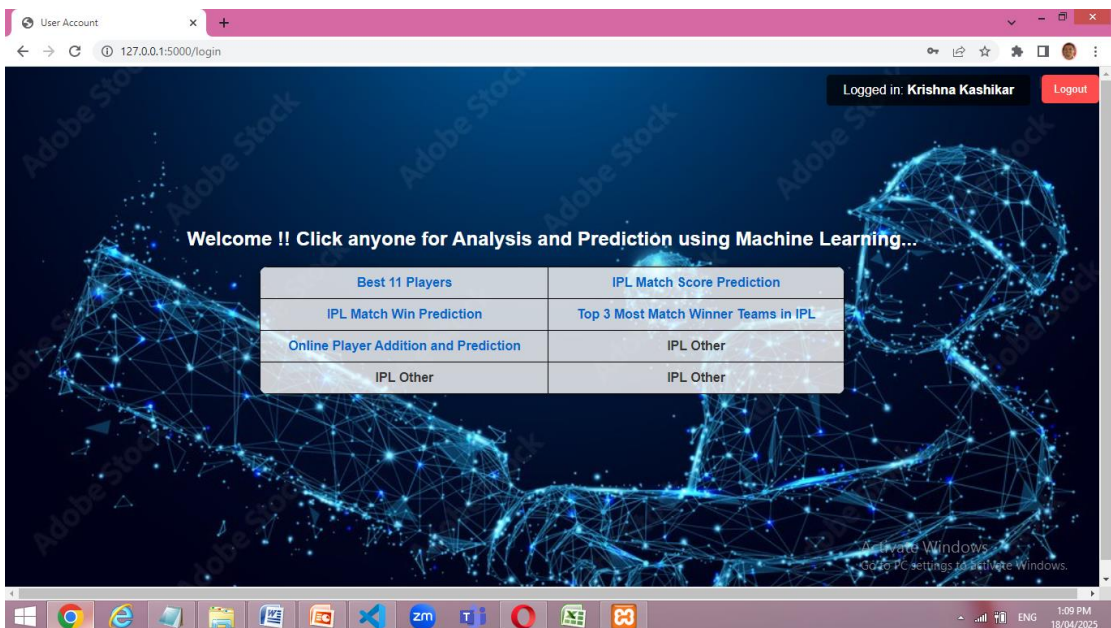
CHAPTER 5
RESULT
AND
DISCUSSION

5 RESULT AND DISCUSSION

AUTHENTICATION: Users begin by selecting their role and registering using their email and password.

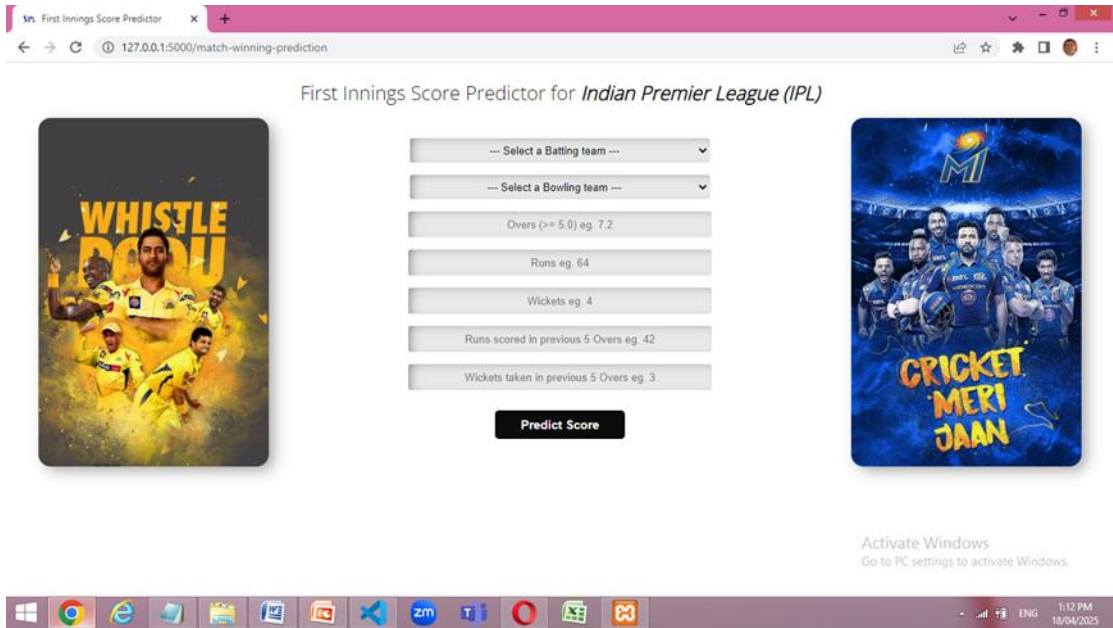


Screenshot 5.1 Authentication page

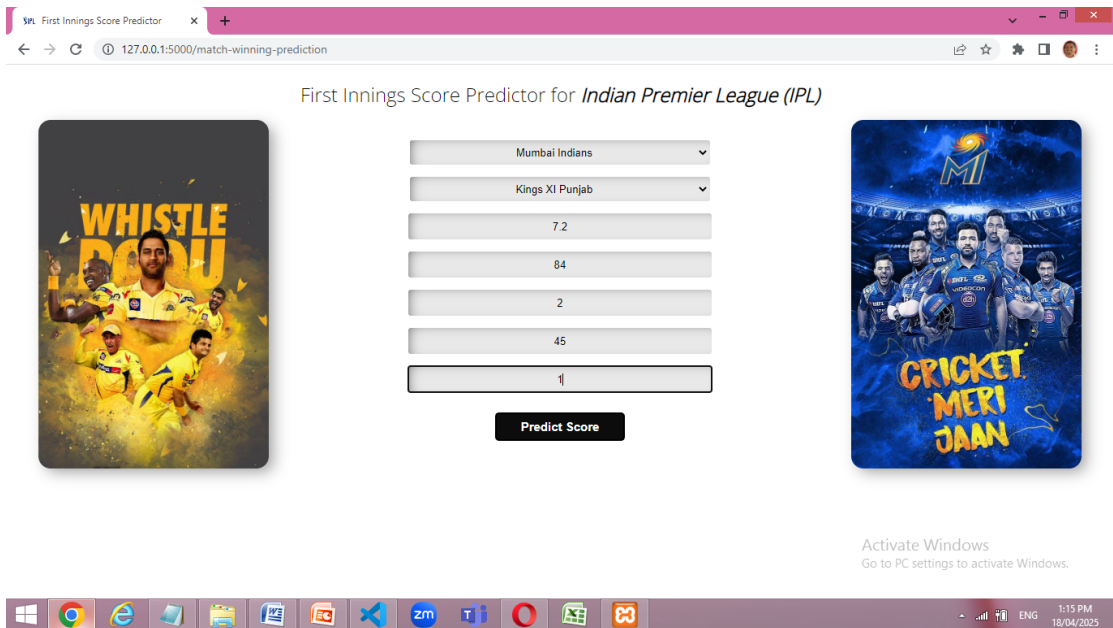


Screenshot 5.2 Home Screen page

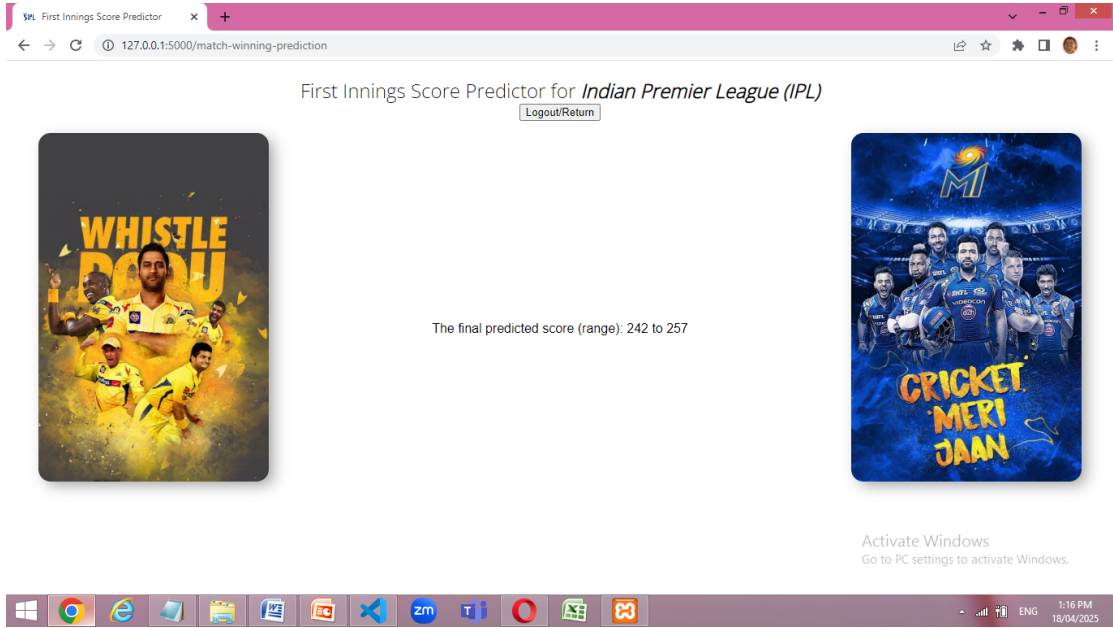
RESULT: The result for First Innings Score Prediction



Screenshot 5.3 First Inning Score Prediction1

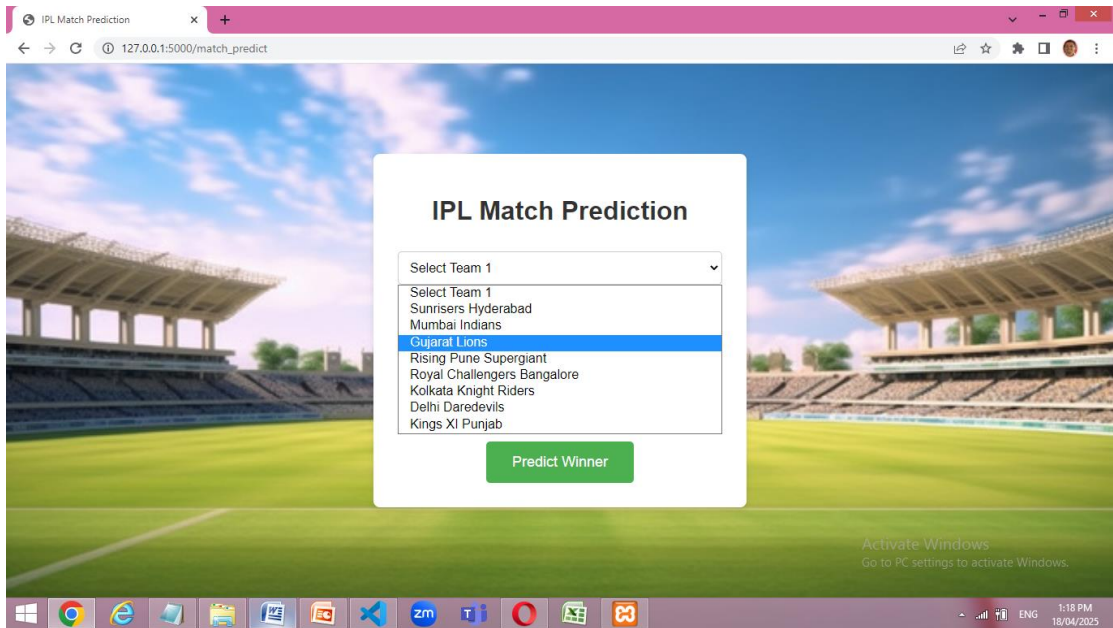


Screenshot 5.4 First Inning Score Prediction2

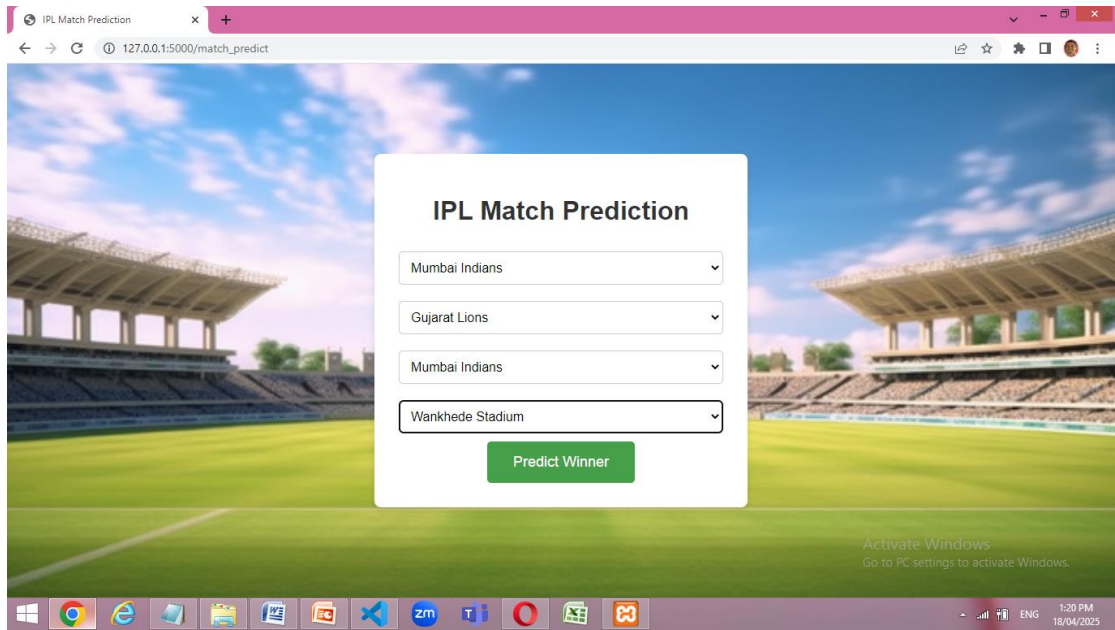


Screenshot 5.5 First Inning Score Prediction3

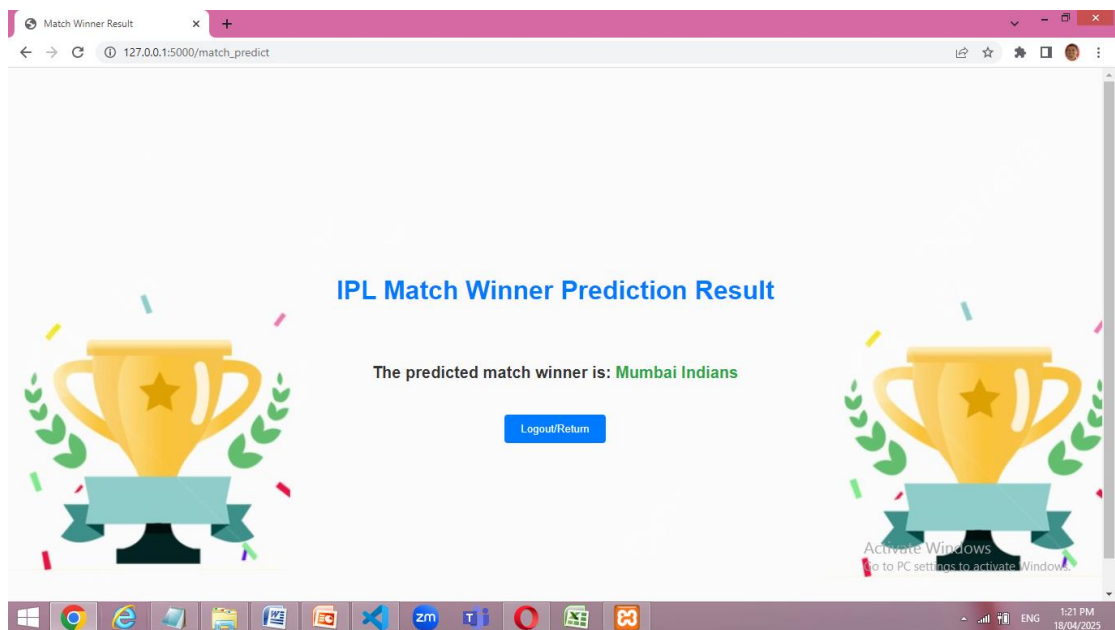
RESULT: For IPL Match win prediction



Screenshot 5.6 IPL Match win Prediction1

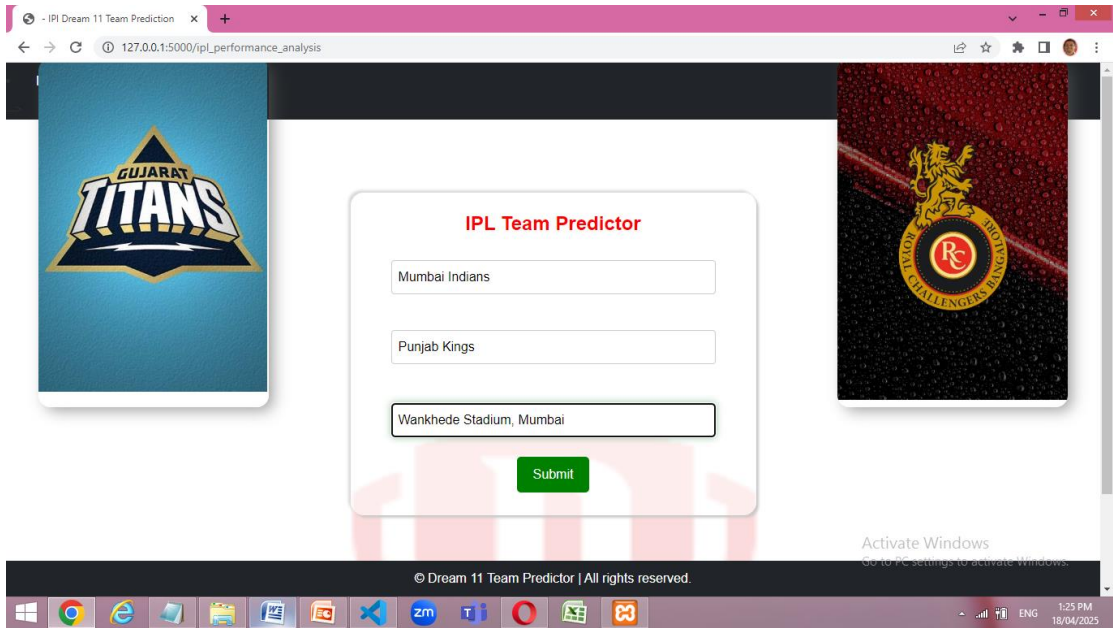


Screenshot 5.7 IPL Match win Prediction2

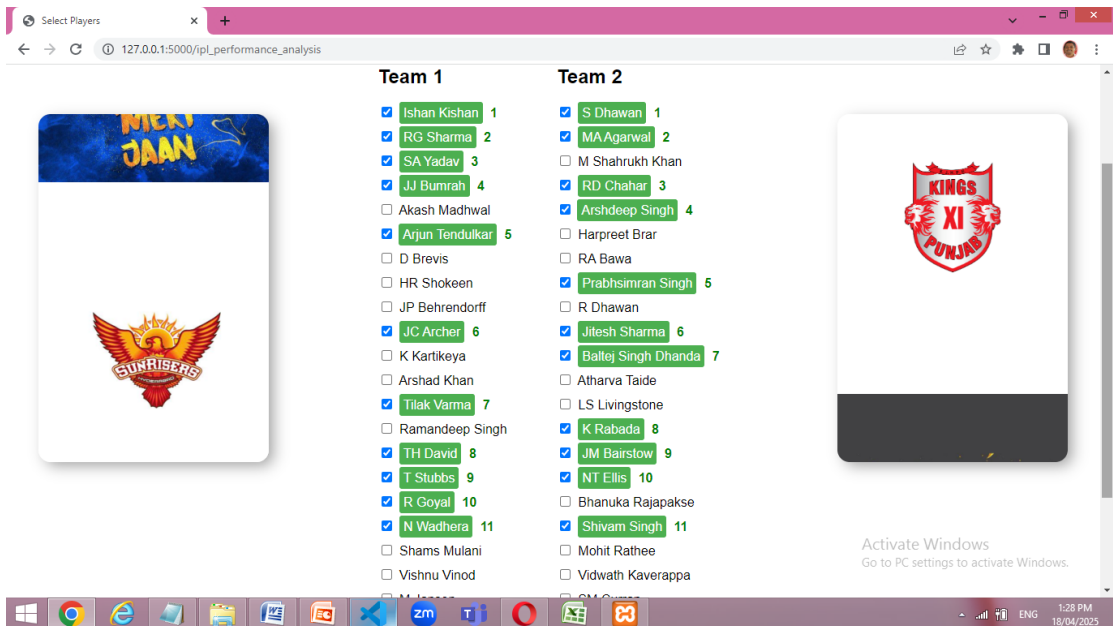


Screenshot 5.8 IPL Match win Prediction3

RESULT: For IPL Team Prediction



Screenshot 5.9 IPL Team Prediction1



Screenshot 5.10 IPL Team Prediction2

The screenshot displays a web application interface for IPL team prediction. On the left is the Rajasthan Royals logo, and on the right is the Delhi Capitals logo. In the center, a table titled 'Final Predicted Team' lists 11 players. Below the table is a 'Logout/Return' button. The browser's address bar shows the URL '127.0.0.1:5000/ipl_performance_analysis'. The Windows taskbar at the bottom shows various application icons and the system tray with the date '18/04/2025' and time '1:31 PM'.

Sr	Players
0	SA Yadav
1	JJ Bumrah
2	RG Sharma
3	Ishan Kishan
4	S Dhawan
6	JC Archer
6	TH David
7	Tilak Varma
8	RD Chahar
9	Arshdeep Singh
10	K Rabada

Screenshot 5.11 IPL Team Prediction3



Model Accuracy: 0.5373

Classification Report:

Class	Precision	Recall	F1-score	Support
0	0.7	0.85	0.77	27.0
1	0.5	0.29	0.36	7.0
2	0.56	0.71	0.62	7.0
3	0.0	0.0	0.0	5.0
4	0.0	0.0	0.0	3.0
5	0.67	1.0	0.8	4.0
6	0.29	0.56	0.38	9.0
8	0.64	0.62	0.63	29.0
9	0.5	1.0	0.67	1.0
10	0.56	0.7	0.62	27.0
11	0.0	0.0	0.0	5.0
12	0.0	0.0	0.0	3.0
13	0.5	0.5	0.5	26.0
14	0.0	0.0	0.0	1.0
15	0.0	0.0	0.0	1.0
16	0.78	0.24	0.37	29.0
17	0.73	0.65	0.69	17.0

Screenshot 5.12 Model Accuracy

Mean Squared Error (MSE): 225.67
 Root Mean Squared Error (RMSE): 15.02
 R² Score: 0.92
 Confusion Matrix: $\begin{bmatrix} 2 & 1 \\ 0 & 1 \end{bmatrix}$

CHAPTER 6
CONCLUSION

6 CONCLUSION

6.1 CONCLUSION

This project successfully demonstrates the application of machine learning in the domain of fantasy sports, specifically for predicting the best XI players in Dream11 cricket matches. By leveraging algorithms such as Random Forest and Decision Tree, the system is able to analyze historical player performance data and provide intelligent, data-driven recommendations for team selection.

The proposed system not only enhances the accuracy of player selection but also improves the overall user experience by automating a typically time-consuming process. It enables users to make more informed decisions, thereby increasing their chances of success in fantasy leagues. The project also incorporates a secure login and registration feature, ensuring user data privacy and a personalized experience.

Through thorough evaluation and comparison of multiple models, the project identifies the most effective approach for fantasy team prediction. The integration of data science and sports analytics in this project highlights the growing potential of machine learning in real-world applications.

6.2 FUTURE SCOPE

This project can be extended by integrating real-time data streams, such as live match data, to enhance prediction accuracy. The same method can be used in order to predict the players from different games like Football, Hockey and Pro-Kabaddi where player selection is very much important. Deep learning techniques like CNN can be explored for modelling sequential or visual data, respectively. The model can also be expanded to include psychological or behavioural factors, such as player pressure handling or consistency under stress. Additionally, incorporating domain-specific rules and constraints for other sports will allow seamless adaptation of the model beyond cricket. Developing a mobile-friendly interface and integrating the system with fantasy league platforms can also broaden its usability for fans and analysts alike.

REFERENCES

REFERENCES

- [1] M. K. Mahbub, M. A. Miah Sikder, M. S. Islam, S. Sorna, S. Hossain, and M. Biswas, "Best Eleven Forecast for Bangladesh Cricket Team with Machine Learning Techniques," 2021 5th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), 2021, pp. 1-6, doi:10.1109/ICEEICT53905.2021.9667862
- [2] A. Basit, M. B. Alvi, F. H. Jaskani, M. Alvi, K. H. Memon, and R. A. Shah, "ICC T20 Cricket World Cup2020 Winner Prediction Using Machine Learning Techniques," 2020 IEEE 23rd International MultitopicConference (INMIC), 2020, pp. 1-6, doi: 10.1109/INMIC50486.2020.9318077
- [3] Saurav Singla , Swapna Samir Shukla "Integer Optimization for Dream 11 Cricket Team Selection" International Journal of Computer Sciences and Engineering · November 2020 DOI: 10.26438/ijcse/v8i11.16
- [4] N. M. Patil, B. H. Sequeira, N. N. Gonsalves, and A. A. Singh, "Cricket Team Prediction Using Machine Learning Techniques,"
- [5] Harshit Barot, Arya Kothari, Pramod Bide, Bhavya Ahir, Romit Kankaria "Analysis and Prediction for the Indian Premier League"2020 International Conference for Emerging Technology (INCET) Belgaum, India. Jun 5-7, 2020
- [6] Sri Ram Teja, T. Pavan Kalyan, V. Akhil Kumar Reddy "Cricket Player Selection using Machine Learning" International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249–8958 (Online), Volume-9 Issue-5, June 2020
- [7] S. BosuBabu, V. Vivek, D. M. T. Kumar, K. Prathyusha, and G. P. Teja, "Predicting football player's position," Int. Res. J. Mod. Eng. Technol. Sci., vol. 4, no. 5, May 2022.
- [8] M. S., V. Bhagat, and P. R. M., "Prediction of football players' performance using machine learning and deep learning algorithms," in 2021 2nd International Conference for Emerging Technology (INCET), 2021, pp. 1–6, doi: 10.1109/INCET51464.2021.9456424
- [9] Shrinivas Biradar, Kavya M, Mahima K. N, Manasa Roopini N, and Meghana M, "Pro Kabaddi winner prediction using logistic regression," Int. J. Adv. Res. Sci. Commun. Technol. (IJARSCT), vol.3, no. 10, May 2023.
- [10] M. K. Mahbub, M. A. M. Sikder, M. S. Islam, S. Sorna, S. Hossain, and M. Biswas, "Best eleven forecasts for Bangladesh cricket team with machine learning

- techniques,” in 2021 5th International Conference on Electrical Engineering and Information & Communication Technology (ICEEICT), 2021, pp. 1–6, doi: 10.1109/ICEEICT53905.2021.9667862.
- [11] A. Basit, M. B. Alvi, F. H. Jaskani, M. Alvi, K. H. Memon, and R. A. Shah, “ICC T20 Cricket World Cup 2020 winner prediction using machine learning techniques,” in 2020 IEEE 23rd International Multitopic Conference (INMIC), 2020, pp. 1–6, doi: 10.1109/INMIC50486.2020.9318077.
- [12] S. Singla and S. S. Shukla, “Integer optimization for Dream 11 cricket team selection,” *Int. J. Comput. Sci. Eng.*, vol. 8, no. 11, Nov. 2020, doi: 10.26438/ijcse/v8i11.16.
- [13] N. M. Patil, B. H. Sequeira, N. N. Gonsalves, and A. A. Singh, "Cricket Team Prediction Using Machine Learning Techniques,"
- [14] M. Ishi, J. Patil, and V. Patil, “An efficient team prediction for one-day international matches using a hybrid approach of CS-PSO and machine learning algorithms,” *Array*, vol. 14, p. 100144, Apr. 2022, doi: 10.1016/j.array.2022.100144.
- [15] B. Sakthivel, A. A. Abinaya, J. R. M. Juraidha, M. T. Madhuvarshini, and T. B. Tamilsudar, “Cricket team prediction using machine learning,” *Int. J. Curr. Eng. Sci. Res. (IJCESR)*, vol. 10, no. 4, pp. 102–106, 2023, doi: 10.21276/ijcesr.2023.10.4.10.
- [16] N. Dhonge, S. Dhole, N. Wavre, M. Pardakhe, and A. Nagarale, “IPL cricket score and winning prediction using machine learning techniques,” *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 3, no. 5, pp. 1723–1728, May 2021.
- [17] S. K. Gupta and S. K. Sharma, “Analyzing and predicting the performance of players using machine learning,” in 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Coimbatore, India, 2023, pp. 1–5.
- [18] S. K. Gupta and S. K. Sharma, “PlayerRank: Leveraging learning-to-rank AI for player positioning in cricket,” in 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Coimbatore, India, 2023, pp. 1–5.
- [19] S. K. Gupta and S. K. Sharma, “Optimal cricket team selection using machine learning,” in 2023 International Conference on Artificial Intelligence and Smart Communication (AISC), Coimbatore, India, 2023, pp. 1–5. Nour El Madhoun, Badis Hammi, “Blockchain Technology in the Healthcare Sector: Overview and Security Analysis,” *IEEE 14th Annual Computing and Communication Workshop and Conference (CCWC)*, Jan. 2024, Las Vegas, NV,

DISSEMINATION OF WORK

DISSEMINATION OF WORK



International Journal of Engineering Technology and Management Sciences

Website: ijetms.in Issue: 2 Volume No.9 March - April – 2025

DOI:10.46647/ijetms.2025.v09i02.041 ISSN: 2581-4621

Performance Analysis and Optimization of Player Selection Using Machine Learning Algorithms

Prof. Ms. Pooja Pohare¹, Krishna M. Kashikar², Gaurav B. Wankhade³, Aniket V. Rajguru⁴, Prashik T. Ingle⁵

¹Asst. prof. Dept. of CSE, Shri Sant Gajanan Maharaj College of Engineering, Shegaon-444203, India

²UG Scholar, Dept. of CSE, Shri Sant Gajanan Maharaj College of Engineering, Shegaon-444203, India

³UG Scholar, Dept. of CSE, Shri Sant Gajanan Maharaj College of Engineering, Shegaon-444203, India

⁴UG Scholar, Dept. of CSE, Shri Sant Gajanan Maharaj College of Engineering, Shegaon-444203, India

⁵UG Scholar, Dept. of CSE, Shri Sant Gajanan Maharaj College of Engineering, Shegaon-444203, India

Abstract

Games like Cricket and Football are very much popular in the world. In India IPL is very much popular, in the similar way games like Pro-Kabaddi, Hockey are also become very popular. Player selection is one the difficult and crucial task in such sports, where the performance of the team is heavily depends on the selection of right players. This research aims to develop an application that employs various algorithms which involves RF, KNN and LR, in order to predict the optimal player line-up based on historical performance data. The study also emphasizes the application of optimization techniques like hyper parameter tuning and cross validation to enhance the accuracy for the prediction of the models. The performance of these models is evaluated using metrics such as Accuracy, Precision, Recall and F1-Score, with results compared to assess the most effective approach for player selection in each sport depending on the selected dataset under one roof.

The results of the study show that all three Machine Learning models – RF, KNN, and LR – can predict player selection with varying degrees of accuracy. After optimization with Grid Search method, Random Forest emerged as the most accurate model for football and cricket, while KNN performed well in Pro-Kabaddi. Expanding beyond traditional Machine learning models to deep learning techniques like Neural Networks could capture more patterns in player performance data.

Keywords: Grid Search Method; KNN; Logistic Regression; Machine Learning; Random Forest

1. INTRODUCTION

An easy Machine learning algorithm, such as Random Forest, K-Nearest Neighbours (KNN), and Logistic Regression, have emerged as powerful tools for predicting player selection in sports analytics. These algorithms leverage historical performance data to forecast the best player line-up for upcoming matches in various sports such as cricket, hockey, Pro-Kabaddi, and football. The ability to analyse vast amounts of player data and make accurate predictions makes these algorithms highly valuable in optimizing team performance. This research focuses on applying these algorithms to predict player selection, with an emphasis on optimizing the models through Grid Search to improve prediction accuracy. By fine-tuning hyper parameters, the research aims to develop an application that can be used in real-time by sports teams to make informed decisions about player selection.

14% Overall Similarity

The combined total of all matches, including overlapping sources, for each database.

Filtered from the Report

- Bibliography

Match Groups

- 33 Not Cited or Quoted 14%**
Matches with neither in-text citation nor quotation marks
- 0 Missing Quotations 0%**
Matches that are still very similar to source material
- 0 Missing Citation 0%**
Matches that have quotation marks, but no in-text citation
- 0 Cited and Quoted 0%**
Matches with in-text citation present, but no quotation marks

Top Sources

- 7% **Internet sources**
- 12% **Publications**
- 4% **Submitted works (Student Papers)**

Integrity Flags

0 Integrity Flags for Review

No suspicious text manipulations found.

Our system's algorithms look deeply at a document for any inconsistencies that would set it apart from a normal submission. If we notice something strange, we flag it for you to review.

A Flag is not necessarily an indicator of a problem. However, we'd recommend you focus your attention there for further review.



CERTIFICATE OF PUBLICATION

This is Awarded to:

Gaurav B. Wankhade

For the paper titled

Performance Analysis and Optimization of Player Selection Using Machine Learning Algorithms

Published in International Journal of Engineering Technology and Management Sciences (IJETMS)

Vol-09 Issue-02 March -April 2025.

DOI: 10.46647/ijetms.2025.v09i02.041 ISSN: 2581-4621

B. Mahesh

Editor-in-Chief







GROUP MEMEBERS

Name- Gaurav B. Wankhade

Address- At. Khedegaon,
Post. Momindabad,

Tq. Nandura, Dist. Buldhana

Mobile No.- 9322107472

Email -
wankhadegaurav71@gmail.com



Name- Krishna Manoj Kashikar

Address- Block no.10,
Tulsidas bhavan, SSGMCE
campus shegaon-444203

Mobile No.- 9175565031

Email-
krishna.kashikar27@gmail.com



Name- Aniket Vishnu Rajguru

Address- Sant Krupa,
Gurudev Nagar,
Digras-445203

Mobile No.-
8208257318

Email-
aniketrajguru2901@gmail.com



Name- Prashik Tejrao Ingle

Address- At. Post -
Kumbefal, Tq-
Khamgaon, Dist -
Buldhana, Pin Code -
444 306

Mobile No.-
9359504931

Email-
prashikingle200@gmail.com



Name- Jivan S. Thorat

Address- At. Patni, Near
Nanduikeshwar Temple, Tq.
Dist. Washim, Pin Code-
444 305

Mobile No.- 7350159072

Email-
jivanthorat7350@gmail.com



