

A PROJECT REPORT ON

“Analysis And Prediction Of Cricket Match Outcome Using Machine Learning ”

Submitted to the Savitribai Phule Pune University

In partial fulfillment for the award of the Degree

OF

BACHELOR OF ENGINEERING

IN

INFORMATION TECHNOLOGY

BY

Suraj Sonawane

B150238627

Sagar Wankhade

B150238640

Rohit Shinde

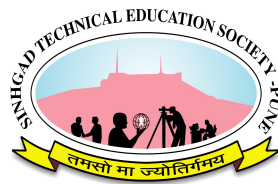
B150238619

Mandar Tannu

B150238632

Under the guidance of

Mr. C.D.KOKANE



Sinhgad Institutes

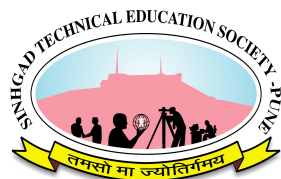
Department of Information Technology

STES's Sinhgad College of Engineering

Vadgaon (Bk.), Off. Sinhgad Road, Pune 411041.

Semester-VIII, Final Year Engineering

2018-2019



Sinhgad Institutes

CERTIFICATE

This is to certify that the project report entitles

**“Analysis and Prediction of Cricket Match Outcome Using
Machine Learning”**

Submitted By

**Suraj Sonawane
Sagar Wankhade
Rohit Shinde
Mandar Tannu**

**B150238627
B150238640
B150238619
B150238632**

are bonafide student group of this institute and the work has been carried out by him/her under the supervision of Mr. C. D. Kokane and it is approved for the partial fulfillment of the requirement of Savitribai Phule Pune University, for the award of the degree of Bachelor of Engineering (Information Technology).

**Mr. C. D. KOKANE
(Project Guide)**

**Prof. G. R. Pathak
(HOD I.T.)**

**Dr. S. D. Lokhande
(Principal)**

This Project Based Seminar report has been examined by us as per the Savitribai Phule Pune University, Pune requirements at STES's Sinhgad College of Engineering, Pune-411041 on

Internal Examiner

External Examiner

ACKNOWLEDGEMENT

We are highly indebted to our guide **Mr. C.D.Kokane** for his guidance and constant supervision as well as for providing necessary information regarding the Project report and also for his support in completing the Project report. We would like to express our special gratitude and thanks to Staff Members of department of Information Technology for giving us such attention and time.

This acknowledgment would be incomplete without expressing our thanks to **Prof. G. R. Pathak**, Head of the Department (Information Technology) for his support during the work.

We would like to extend our heartfelt gratitude to Principal, **Dr. S. D. Lokhande** who provided a lot of valuable support, mostly being behind the veils of college bureaucracy.

We would also like to express our gratitude towards our parents and friends for their kind co-operation and encouragement which help us in completion of this Project. Our thanks and appreciations also go to our colleague in developing the Project report and people who have willingly helped us out with their abilities.

SURAJ SONAWANE
SAGAR WANKHADE
ROHIT SHINDE
MANDAR TANNU
(B.E. I.T)

ABSTRACT

Cricket is one of the most popular team games in the world. World cup is the biggest cricket tournament. The ICC Cricket World Cup is the international championship of One Day International (ODI) cricket. The event is organized by the sport's governing body, the International Cricket Council (ICC), every four years, with preliminary qualification rounds leading up to a finals tournament. The tournament is one of the world's most viewed sporting events and is considered the "flagship event of the international cricket calendar" by the ICC. Essentially, it is difficult to find optimal set of attributes which would lead to assemble a team with highest chance of winning. Several efforts have been made to analyse factors affecting outcomes of a match without much success. To overcome these issues fundamental concepts of analytics and predictive modeling to ICC cricket matches can be used to get meaningful information and predictions. Java is a general-purpose computer-programming language that is concurrent, class-based, object-oriented, and specifically designed to have as few implementation dependencies as possible. To design web pages, we are using jsp servlet and for dynamic web page calling we are using apache tomcat server, For handling database in the backend MySQL has been used.

Contents

Certificate	ii
Acknowledgement	iii
Abstract	iv
1 Introduction	1
1.1 Motivation	2
1.2 Problem Definition	2
2 Literature Survey	3
3 Software requirement specification	7
3.1 Introduction	7
3.1.1 Purpose	7
3.1.2 Document Conventions	7
3.1.3 Intended Audience And Reading Suggestions	7
3.1.4 Project Scope	7
3.1.5 References	8
3.2 Overall Description	8
3.2.1 Product Perspective	8
3.2.2 Product Features	8
3.2.3 User Class and Characteristics	10

3.2.4	Operating Environment	10
3.3	System Features	10
3.4	External Interface Requirements	11
3.4.1	User Interfaces	11
3.4.2	Hardware Interfaces	11
3.4.3	Software Interfaces	11
3.4.4	Communication Interfaces	12
3.5	Nonfunctional Requirements	12
3.5.1	Performance Requirements	12
3.5.2	Safety Requirements	12
3.5.3	Security Requirements	12
3.5.4	Software Quality Attributes	12
4	System Design	14
4.1	System Architecture:	14
4.2	Dataflow Diagram:	14
4.3	UML Diagrams:	16
4.3.1	USE CASE Diagram:	16
4.3.2	Class Diagram:	17
4.3.3	Sequence Diagram	18
5	Project Implementation	19

5.1	Introduction	19
5.2	Algorithm Used	20
6	Software Testing	23
6.1	Type of Testing Used	23
6.2	Test Cases and Test Results	24
7	Results	26
7.1	Screen shots	26
8	Deployment and Maintenance	39
8.1	Operational Requirements to run the System	39
8.1.1	Installing Software	39
8.1.2	Uninstalling Software	40
8.2	User help	41
9	Conclusion and Future Scope	42
9.1	Conclusion and Future Scope	42
10	References	43
A	Testing Of Project Problem Statement Using Generated Test Data:	44
B	Project Planner	46

List of Figures

1	ER Diagram	9
2	System Architecture:	14
3	DFD 0	14
4	DFD 1	15
5	DFD 2	15
6	Use Case	16
7	Class Diagram	17
8	Sequence Diagram	18
9	Formula	21
10	Login	26
11	User registration	27
12	Player Details	28
13	Batsmen Details	29
14	Bowler Details	30
15	All Rounder(i)	31
16	All Rounder(ii)	32
17	Insert Team	33
18	AssignPlayer	34
19	Predict UI	35
20	Result	36

21	Probable Playing 11	37
22	Database	38

List of Tables

1	Literature Survey	6
2	Software Interfaces	11
3	Test Cases	25

List of Abbreviations

- DB - Database
- NB - Naïve Bayes
- ER - Entity Relationship Diagram

1 Introduction

Now-a-days sports have much importance in our life as well as on national and international level. Cricket is a sport with multiple formats, different playing standards and varying durations. ODI is one of the three current forms of Cricket which is recognized by the International Cricket Council (ICC). Because of the excitement it generates, ODI cricket has become such a huge success. There are many annual tournaments conducted at both domestic and International level.

The format of the Cricket World Cup has changed greatly over the course of its history. Each of the first four tournaments was played by eight teams, divided into two groups of four. The competition consisted of two stages, a group stage and a knock-out stage. The four teams in each group played each other in the round-robin group stage, with the top two teams in each group progressing to the semi-finals. The winners of the semi-finals played against each other in the final. With South Africa returning in the fifth tournament in 1992 as a result of the end of the apartheid boycott, ten teams played each other once in the group phase, and the top four teams progressed to the semi-finals. The tournament was further expanded in 1996, with two groups of six teams. The top four teams from each group progressed to quarter-finals and semi-finals.

Estimating cricket match outcome becomes difficult due various varying factors such as no of teams participates in Championship changes, different venues,

In this project, efforts are taken to design a model in which the winning team can be predicted and playing 11 of both teams can be estimated. In the former case Naive Bayes Classification can be used. Unlike the current procedure for predicting the winner, the factors like the venue of the match, the playing 11 of the batting team and past record of that team have been considered in the estimation. These past records have been taken from all the non-curtailed ODI matches played among the top ten international teams.

1.1 Motivation

Cricket is 2nd most popular sport all over the globe. Big data can be used to improve training and understanding competitors, using sport sensors. It is also possible to predict winners in a match using big data analytics. Future performance of players could be predicted as well. Thus, players' value and salary are determined by data collected throughout the season.

1.2 Problem Definition

“To design and implement prediction and analysis of the cricket match outcome like winning team and Playing 11 prior to its commencement using Machine Learning technique.”

2 Literature Survey

Sr. No.	Year	Title	Authors	Description
1	2010	CricAI: A Classification Based Tool to Predict the Outcome in ODI Cricket.	Amal Kalu-arachchi and Aparna S. Varde	In this paper authors have used Bayesian classifiers in machine learning, to predict how these factors affect the outcome of an ODI cricket match. They have developed a software tool called CricAI which gives probability of victory in an ODI cricket match prior to its commencement.
2	2015	Score and Winning Prediction in Cricket through Data Mining	Tejinder Singh, Vishal singla and Parteek Bhatta	In this paper authors proposed two methodology, first it predict the score first inning by considering no. of wickets fallen, venue of the match, batting team. Second it predict the second inning score by same attributes along with target given by opponent team by using Linear regression and naïve Bayes Classifiers.

3	2017	Analysis and Prediction of Sentiments for Cricket Tweets Using Hadoop.	Bharati S. Kannolli and Prabhu R. Bevinmarad	In this paper, authors suggested a Technique to Analyze the outcome of Cricket match. They have proposed an unsupervised method to analyze a data collected from Twitter social media. Experimental analysis was performed using virtual machine CLOUDERA software that run on Hadoop platform. The system was tested both offline and online using large number of tweets.
---	------	--	--	---

4	2017	Identifying the Optimal Set of Attributes that Impose High Impact on the End Results of a Cricket Match Using Machine Learning	Pranavan Somaskandhan, Gihan Wijesinghe, Leshan Bashitha Wijegunawardana, Asitha Bandaranayake and Sampath Deegalla	This paper authors proposed a methodology which is used to find set of attributes that have an impact on the result of a cricket match by using statistical analysis and machine learning models. They have used Extra tree algorithm, naïve Bayes algorithm, support vector machine to find out the best possible attributes which has an impact on cricket match outcome. Further they concluded that SVM has better accuracy.
---	------	--	---	--

5	2017	Cricket Team Prediction with Hadoop: Statistical Modelling Approach	Shubham Agarwal, Lavish Yadav and Shikha Mehta .	Authors predicted suitable Team to be lined for a particular match. Authors used Statistical modelling approach to predict the perfect players for the match to be played. Many Factors were taken into consideration such as Player's Overall Stats, Player Performances with different Teams and the most important Last 5 Performances for selection of players in playing 11 from the Team of 16. Experimental analysis was performed using Hadoop and Hive for Indian players.
---	------	---	--	---

Table 1: Literature Survey

3 Software requirement specification

3.1 Introduction

3.1.1 Purpose

The purpose of this document is to build a system to predict the winner of the ICC world cup match prior to its commencement.

3.1.2 Document Conventions

This document uses the following conventions.

- DB Database
- DDB Distributed Database
- ER Entity Relationship

3.1.3 Intended Audience And Reading Suggestions

This project is a predictive model used to predict the outcomes of cricket match prior to its beginning and it is restricted to world cup matches only. This has been implemented under the guidance of college professors.

3.1.4 Project Scope

The fundamental concepts of analytics and predictive modelling to world cup cricket matches will be applied to get meaningful information and predictions. Teams, matches, and factors affecting outcomes of matches will be analysed. Some factors that affect match outcomes could be venue (stadium), city, toss winner, and toss decision (field/bat). This model is reliable and easy to use. Further, this data can be useful for legal betting like for Dream 11 app.

3.1.5 References

- <https://en.wikipedia.org/wiki/Machinelearning>
- <https://en.wikipedia.org/wiki/NaiveBayesclassifier>
- <https://dzone.com/articles/ipi-cricket-analytics-and-predictive-model>

3.2 Overall Description

3.2.1 Product Perspective

- User Description

It includes username and password. This information is used for login into the system to see the outcomes of cricket match.

- Classifier Description

It contains machine learning technique that is Naïve bays which will handle the independent attributes.

- Analysis Description

It contains several factors affecting the result of a cricket match like ground history, team1, team2, match-winner-count of previous record.

- Dataset Description

It contain contains several attributes like recordid, team1, team2, venue, city.

3.2.2 Product Features

The major feature analysis and prediction of cricket match outcome is shown in below entity – relationship diagram (ER model).

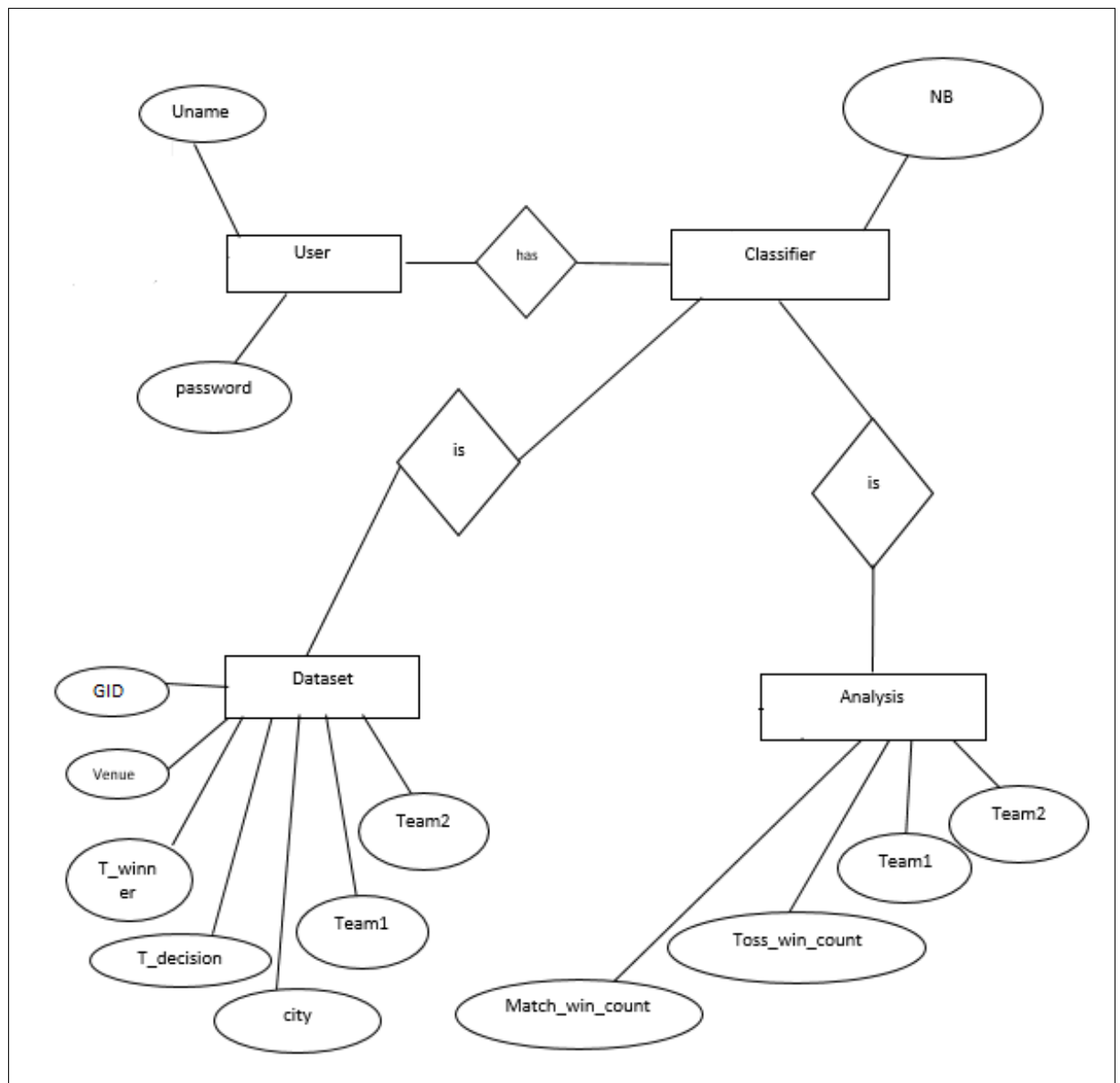


Figure 1: ER Diagram

3.2.3 User Class and Characteristics

User should be able to login to the system and he will be able to know the outcome of a cricket match prior to its commencement. The system has a classifier and by considering previous trained data it will automatically provide the result of match. The system for providing results considers various factors like opposition team, venue, previous records, etc.

3.2.4 Operating Environment

Operating environment for the analysis and prediction of cricket match outcome is as listed below.

- Database: MySQL
- Operating System: Windows
- Server: Apache Tomcat Server
- Language: Java

3.3 System Features

- Gathering statistical data.
- Analysing gathered data.
- Drawing conclusion from analyzed data.
- Using Naïve Bayes classifier, for predicting outcome from statistical data.
- Using the predictive model to predict win/lose.
- And finally system will predict the winner of the match and playing 11 of both the teams.

3.4 External Interface Requirements

3.4.1 User Interfaces

- Front-end software: Java
- Back-end software: MySQL

3.4.2 Hardware Interfaces

Since the application must run over the internet, all the hardware shall require to connect internet will be hardware interface for the system. As for e.g. Modem, WAN – LAN, Ethernet Cross-Cable.

3.4.3 Software Interfaces

Following are the software used for the analysis and prediction of cricket match outcome application.

Software used	Description
Operating system	We have chosen Windows operating system for its best support and user-friendliness.
Database	To save the user data, predictive model algorithms we have chosen MySQL database in Apache Tomcat Server.
Java	To implement the project we have chosen Java language for its more interactive support.

Table 2: Software Interfaces

3.4.4 Communication Interfaces

- Web application run on a web browser.
- Communication standards used is HTTP.

3.5 Nonfunctional Requirements

3.5.1 Performance Requirements

- The performance of the functions and every module must be well.
- The overall performance of the software will enable the users to work efficiently.

3.5.2 Safety Requirements

If there is extensive damage to a wide portion of the database due to catastrophic failure, such as a disk crash, the recovery method restores a past copy of the database that was backed up to archival storage (typically tape) and reconstructs a more current state by reapplying or redoing the operations of committed transactions from the backed up log, up to the time of failure.

3.5.3 Security Requirements

Security systems need database storage just like many other applications. However, the special requirements of the security market mean that vendors must choose their database partner carefully.

3.5.4 Software Quality Attributes

Our software has many quality attribute that are given below :-

- **Adaptability:** This software is adaptable by all users.
- **Availability:** This software is freely available to all users. The availability of the software is easy for everyone.
- **Maintainability:** After the deployment of the project if any error occurs then it can be easily maintained by the software developer.
- **Reliability:** The performance of the software is better which will increase the reliability of the Software.
- **User Friendly:** Since, the software is a GUI application, the output generated is much user friendly in its behaviour.
- **Integrity:** Integrity refers to the extent to which access to software or data by unauthorized persons can be controlled.
- **Security:** Users are authenticated using many security phases so reliable security is provided.
- **Testability:** The software will be tested considering all the aspects such as ground location with different teams.

4 System Design

4.1 System Architecture:

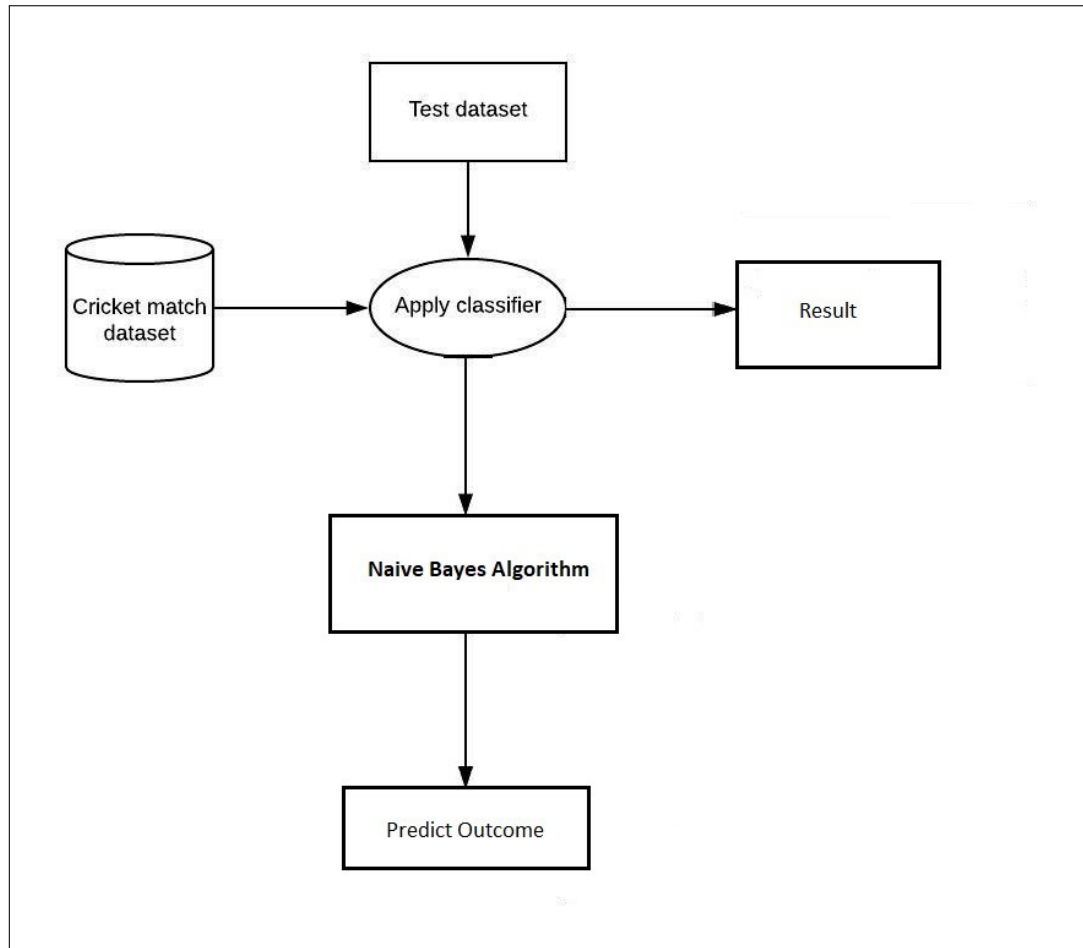


Figure 2: System Architecture:

4.2 Dataflow Diagram:

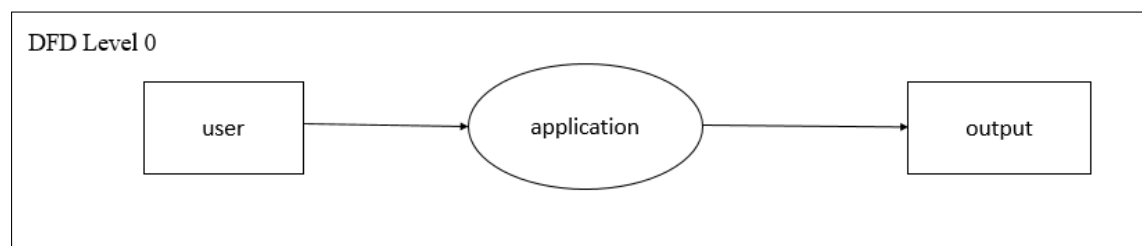


Figure 3: DFD 0

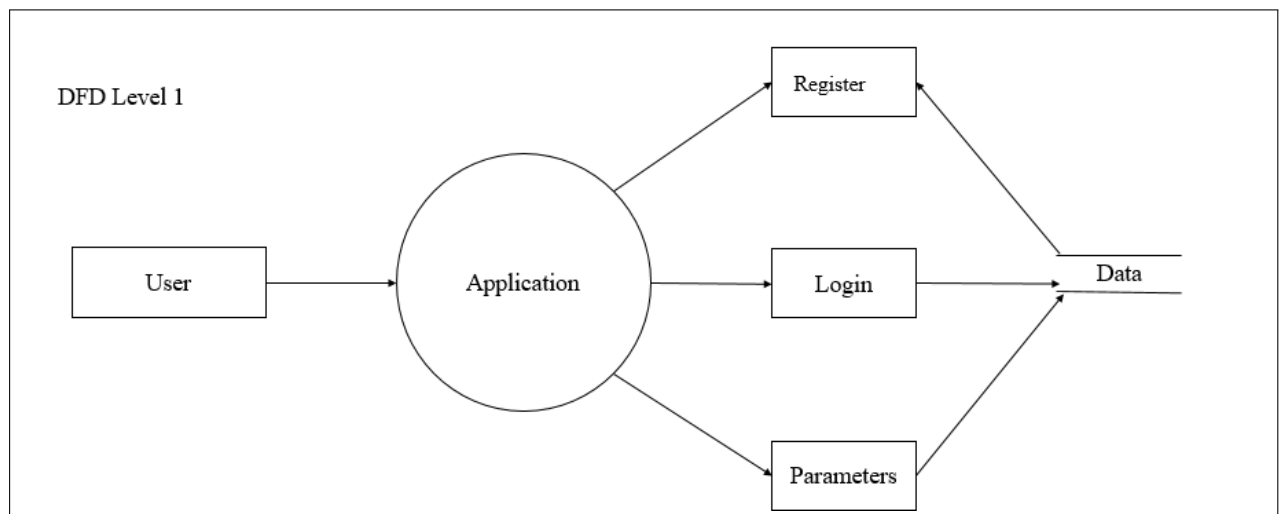


Figure 4: DFD 1

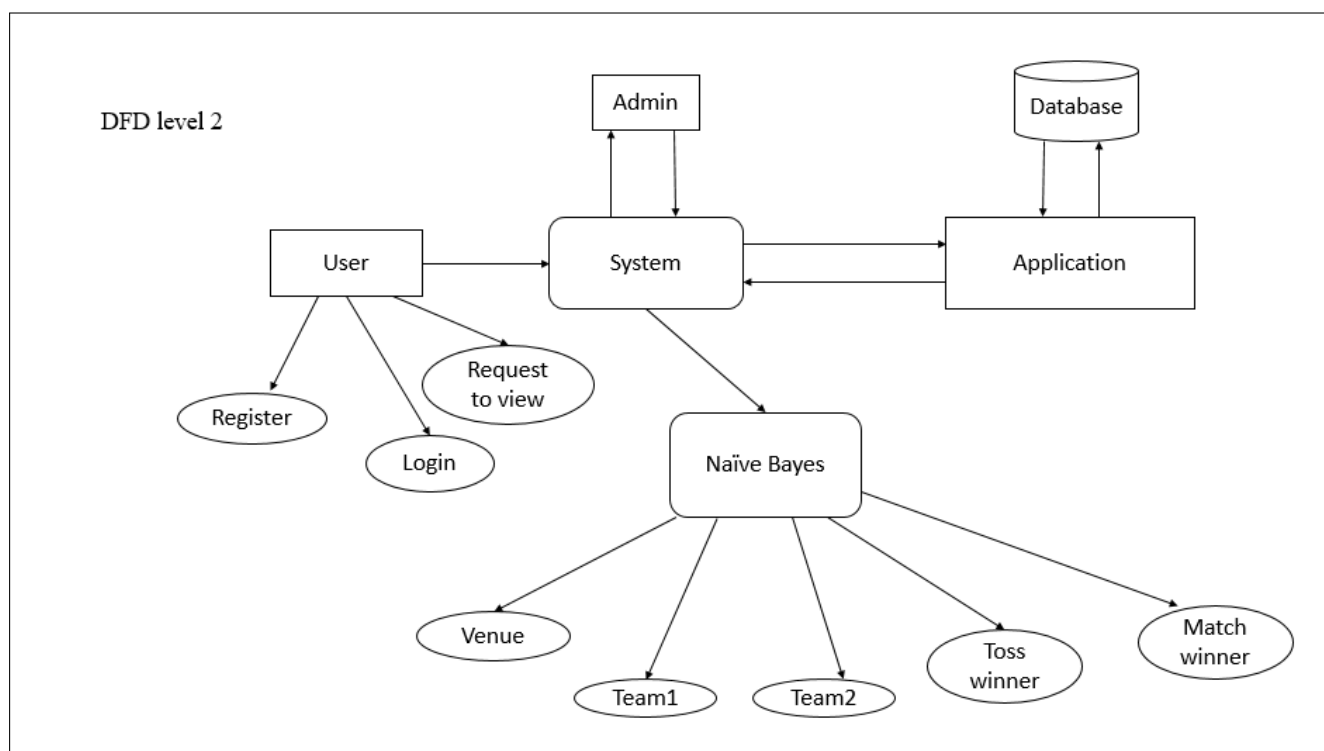


Figure 5: DFD 2

4.3 UML Diagrams:

4.3.1 USE CASE Diagram:

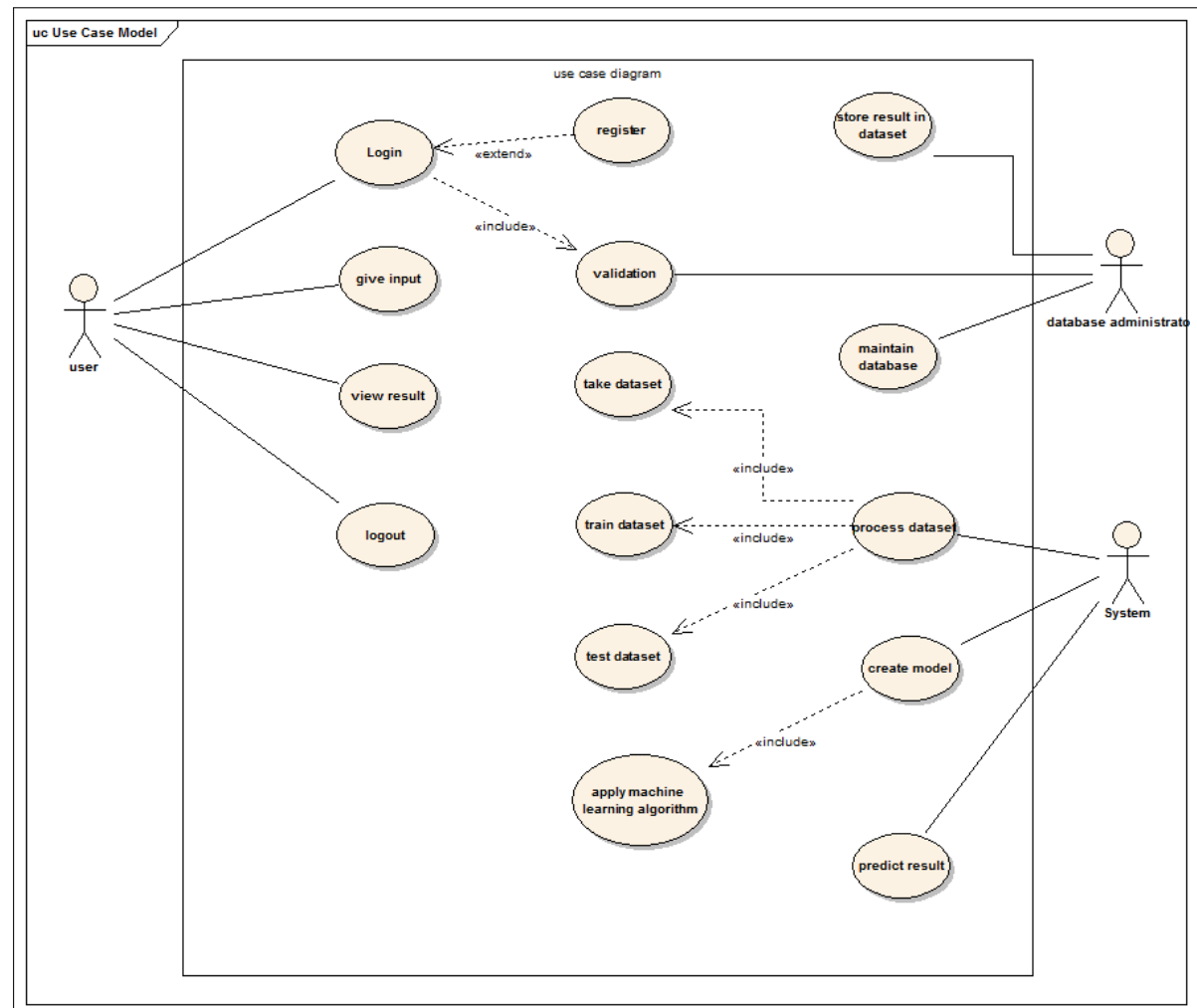


Figure 6: Use Case

4.3.2 Class Diagram:

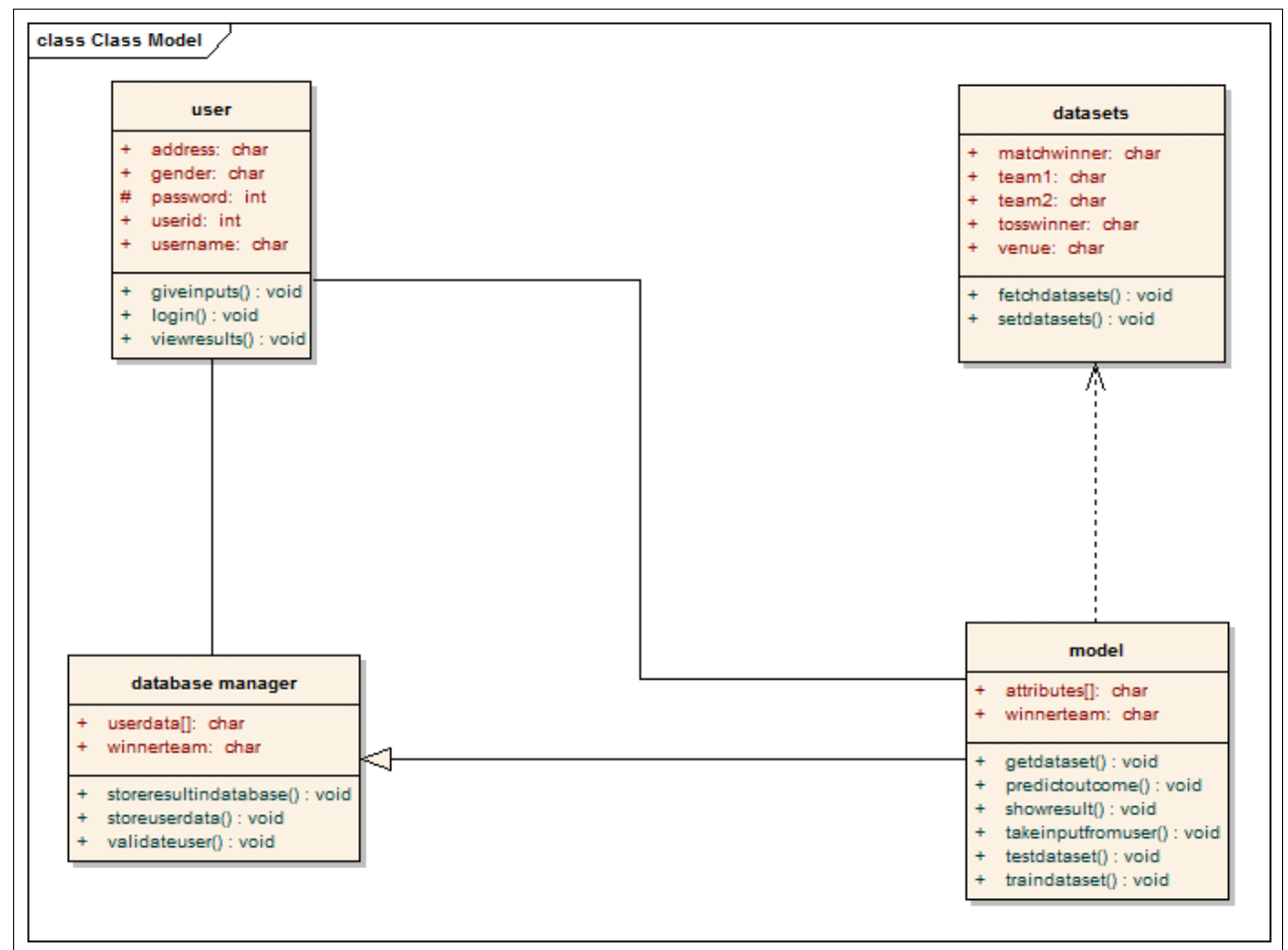


Figure 7: Class Diagram

4.3.3 Sequence Diagram

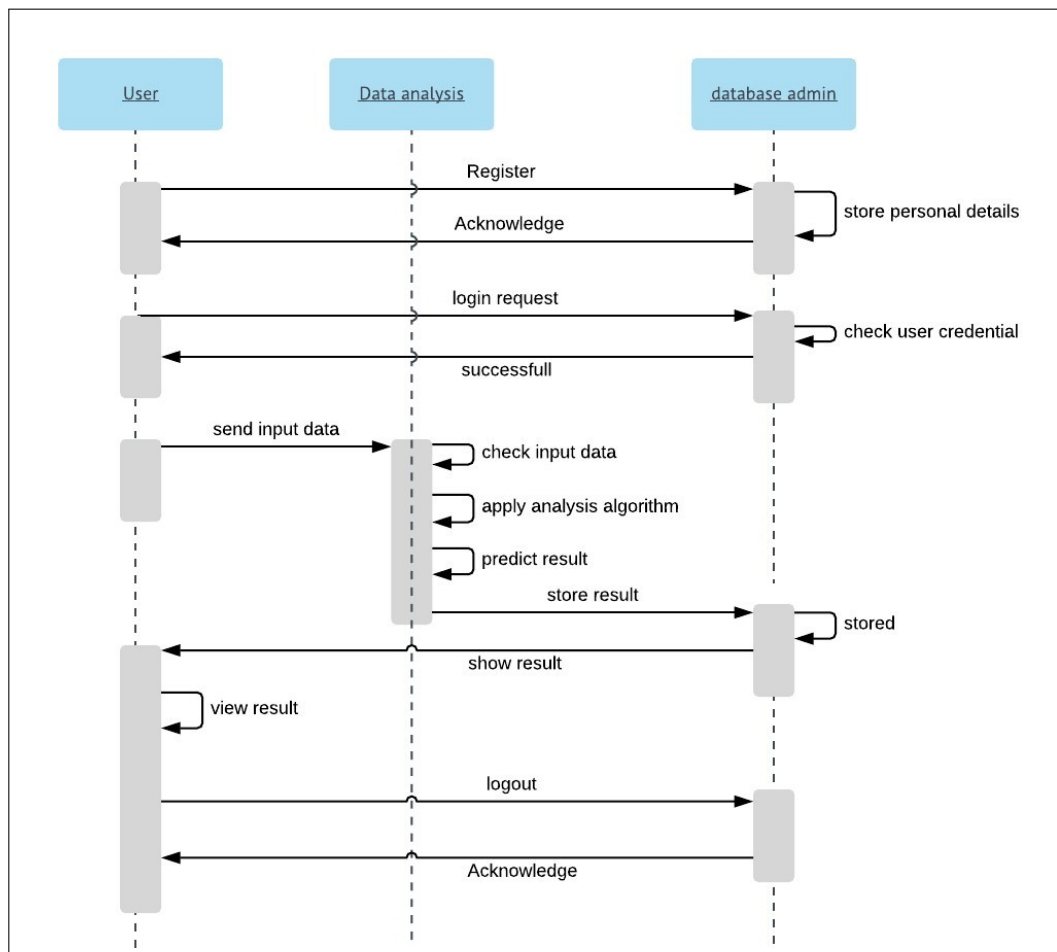


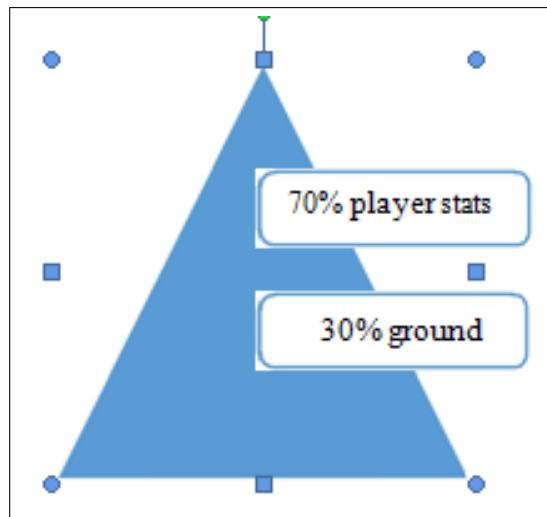
Figure 8: Sequence Diagram

5 Project Implementation

5.1 Introduction

This section proposes an algorithm used to model the batsmen, bowlers and the all-rounders of a team. As cricket requires many manual and on spot decisions, some assumptions are taken. Firstly, it is considered that each and every player is fit for playing. Secondly, it is independent of player's performance in trials (as there is no information available for the same).

For a player to be selected in a team, all factors about cricket must be clear such as categorizing the players into three different groups, Batsmen, Bowler and All Rounder. There are subcategories of Batsmen i.e. Top-Order Batsmen and Middle-Order Batsmen. Subcategories of Bowlers include Spin, Medium-Pace and Fast bowlers. For the Batting Performance, Batting Scores of a player are calculated. Similarly, for the Bowling Performance Bowling Scores of a player are calculated and for the All Rounder's Performance a threshold unit of Batting and Bowling Score is set. Players having scores more than threshold in both comes under the category of All Rounder. For Wining Prediction, weights have been allotted to particular columns as given below:



Player score :

To calculate player score, first data is fetched from database for both teams. At the

beginning all attributes are initiated to zero then each parameter of one team is compared with equivalent parameter of other team. If the one team parameter is greater than other team parameter then we set a integer such as 9 or 10 to the corresponding attributes same method is apply to all 9 attributes. Now team score is calculated by adding all 9 attributes for both teams. At the end total score of both teams are compared.

If(team1marks(battingAvg) > team2marks(battingAvg))

Player1score = 10;

Else

Player2score = 10;

- **Ground score :**

Ground score is significant factor in prediction. Team which have performed well at particular ground has more chances to win the match on same ground.

Ground score is calculated by

If(groundMatches > 0) then

groundUnits = groundWinner / groundMatches

Ground score is added to player score and final score is compared with other team score. Team having higher score is predicted as winning team.

- **Playing 11 :**

On basis of past performance of players, list is created in decreasing order of performance. Top five batsmen are picked from the list and remaining are picked from all-rounders and bowlers.

5.2 Algorithm Used

What is Naive Bayes algorithm?

It is a classification technique based on Bayes' Theorem with an assumption of independence among predictors. In simple terms, a Naive Bayes classifier assumes that

the presence of a particular feature in a class is unrelated to the presence of any other feature. For example, a fruit may be considered to be an apple if it is red, round, and about 3 inches in diameter. Even if these features depend on each other or upon the existence of the other features, all of these properties independently contribute to the probability that this fruit is an apple and that is why it is known as ‘Naive’.

Naive Bayes model is easy to build and particularly useful for very large data sets. Along with simplicity, Naive Bayes is known to outperform even highly sophisticated classification methods. Bayes theorem provides a way of calculating posterior probability $P(c|x)$ from $P(c)$, $P(x)$ and $P(x|c)$. Look at the equation below:

The diagram shows the formula for the posterior probability in Naive Bayes classification. The formula is $P(c|x) = \frac{P(x|c)P(c)}{P(x)}$. Arrows point from labels to the corresponding parts of the formula: 'Likelihood' points to $P(x|c)$, 'Class Prior Probability' points to $P(c)$, 'Posterior Probability' points to $P(c|x)$, and 'Predictor Prior Probability' points to $P(x)$. Below the formula, the expanded version is given: $P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$.

$$P(c|x) = \frac{P(x|c)P(c)}{P(x)}$$

$P(c|X) = P(x_1|c) \times P(x_2|c) \times \dots \times P(x_n|c) \times P(c)$

Figure 9: Formula

Above,

- $P(c|x)$ is the posterior probability of class (c , target) given predictor (x , attributes).
- $P(c)$ is the prior probability of class.
- $P(x|c)$ is the likelihood which is the probability of predictor given class.
- $P(x)$ is the prior probability of predictor.

Let's understand it using an example. Below I have a training data set of weather

and corresponding target variable ‘Play’ (suggesting possibilities of playing). Now, we need to classify whether players will play or not based on weather condition. Let’s follow the below steps to perform it.

Step 1: Convert the data set into a frequency table

Step 2: Create Likelihood table by finding the probabilities like Overcast probability = 0.29 and probability of playing is 0.64.

Step 3: Now, use Naive Bayesian equation to calculate the posterior probability for each class. The class with the highest posterior probability is the outcome of prediction.

6 Software Testing

Software testing is a method for examining the ability of a framework and confirms that it meets its results. It is produced by developers to keep up the quality of programming, software testing still remains a craftsmanship, because of less comprehension of the terms of testing. Fundamental issue in regards to the software testing is from the complication of programming: we can't test entire program with less complexity.. Testing is more than just debugging. The purpose of testing can be quality assurance, verification and validation, or reliability estimation. There are 2 major types of testings:

1. Correctness testing
2. Reliability testing

Software testing has to deal between expenses, time and quality. Presently we will see the test cases for the making website. We execute functional non-functional testing system. Software testing shall design test cases with higher probability of finding mistakes. To meet the correct objective and to give the efficient test cases the maximum numbers of errors should be reported.

6.1 Type of Testing Used

The various types of testing that may be used in the dissertation are as follows:

1. **Smoke testing:** Smoke test is an primary set of tests that decide whether another product assemble is performing alright to acknowledge it for a noteworthy testing effort. It verifies the major functionality at high level. The Smoke test process focuses more on breadth than the depth. If the test is not successful, that test case goes back to developers without testing.
2. **Unit Testing:** Unit testing focuses on the verification effort on the part of system outline the product part or module. The relative complication of tests and revealed bugs is controlled by the restricted scope built up for unit testing. The test goes under the white-box testing, and the progression can be led in parallel for numerous segments.

3. **Integration testing:** Integration testing is used to check whether the small components accurately interact with each other according to the instruction flow.
4. **Validation Testing:** Validation testing can be done by various ways, but a more common definition is that succeeds when Software strategies in a way that can be sensibly expected by the client. Programming approval is acquired through a progression of black box testing that shows with prerequisites. A test plan consists the classes of test derived and a test method describes particular test cases that will be utilized to show conformity with requirements. Both the plan and strategy are build to ensure that all requirements are satisfied, all execution necessities are obtained, all performance requirements are attained, documentation is accurate, and human designed and other necessities are met.
5. **System Testing:** This tests the whole system. It is a functional testing, performed to approve that the application meets user requirement.
6. **Regression Testing:** This is also called as re-testing of a program after the modification. It guarantees that the errors are not coming with result after the changes are done, and that the modified system still meets its results. It is used whenever the software or its environment is changed.
7. **Load Testing:** It includes to test an application when heavy loads of applications is there, so that prediction of application could be observed in real time application. It also determines the point at which the systems response time reduces or fails.

6.2 Test Cases and Test Results

Following table shows the test cases related to the project. Here applied software testing cases which is done by checking each functionality. It is divided into four columns, test case id, test case name, expected result and actual result.

Test Case id	Testcase Name	Expected Result	Actual Result	Status
1	Validate Login	If wrong id and pass-word an error message "Invalid login or pass-word should generated"	same as Expected	pass
2	Text Field Validation	Validate text field an error message "Only text field allowed" should generated	same as Expected	pass
3	Numeric Field Validation	Validate numeric field an error message "Only numbers allowed" should generated	same as Expected	pass

Table 3: Test Cases

7 Results

7.1 Screen shots

Login

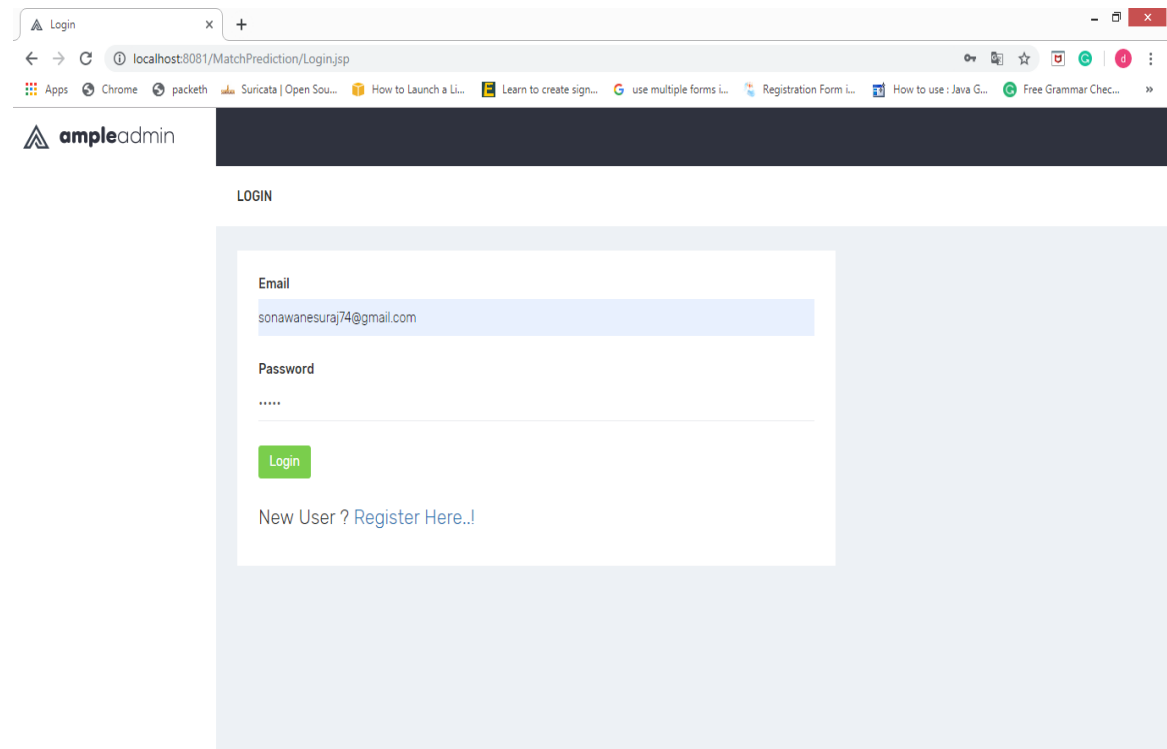


Figure 10: Login

For using this system user have to register himself/herself to the system with login Id and Password as shown in figure below.

The screenshot shows a web browser window with the address bar displaying 'localhost:8081/MatchPrediction/Register.jsp'. The browser's tab is labeled 'Register'. The website's header features the 'ampleadmin' logo. The main content area is titled 'REGISTER' and contains a registration form with the following fields:

- Name:** Sagar
- Email:** wankhadesagar261@gmail.com
- Password:** (masked)
- Mobile:** 8888858565

A green 'Login' button is positioned below the mobile number field.

Figure 11: User registration

Admin can add Player details and the player can be of type batsmen,bowler or allrounder.

The screenshot shows a web browser window with the URL `localhost:8081/MatchPrediction/AddPlayer.jsp`. The page is titled "ADD PLAYER" and features a sidebar with navigation options: Home, Add Ground, Add Team, Add Player (highlighted), and Logout. The main content area is titled "Player Details" and contains three input fields: "Player Name" with the value "Shikhar Dhawan", "Matches Played" with the value "128", and "Player Type" with a dropdown menu showing "Batsmen". A green "Add" button is located at the bottom of the form.

Figure 12: Player Details

If the player is Batsmen

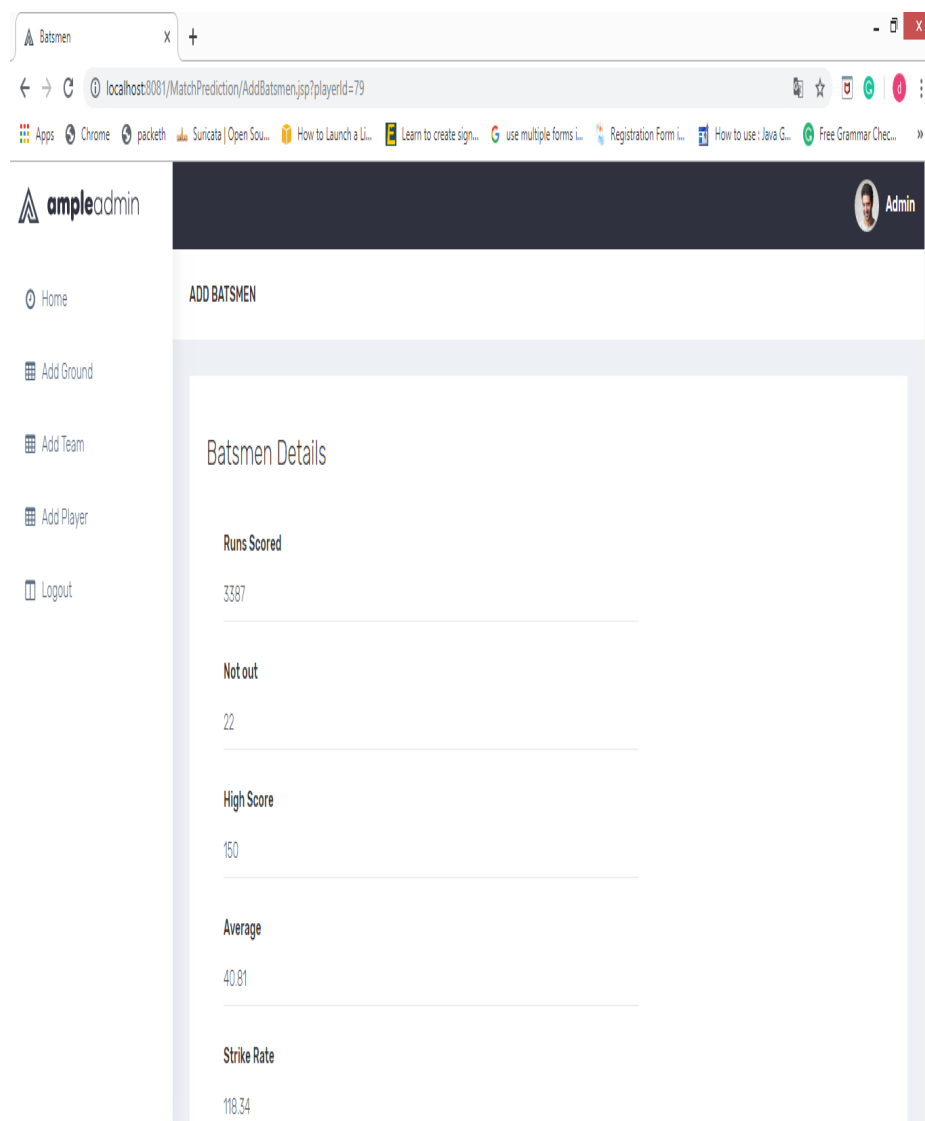


Figure 13: Batsmen Details

If the player is Bowler

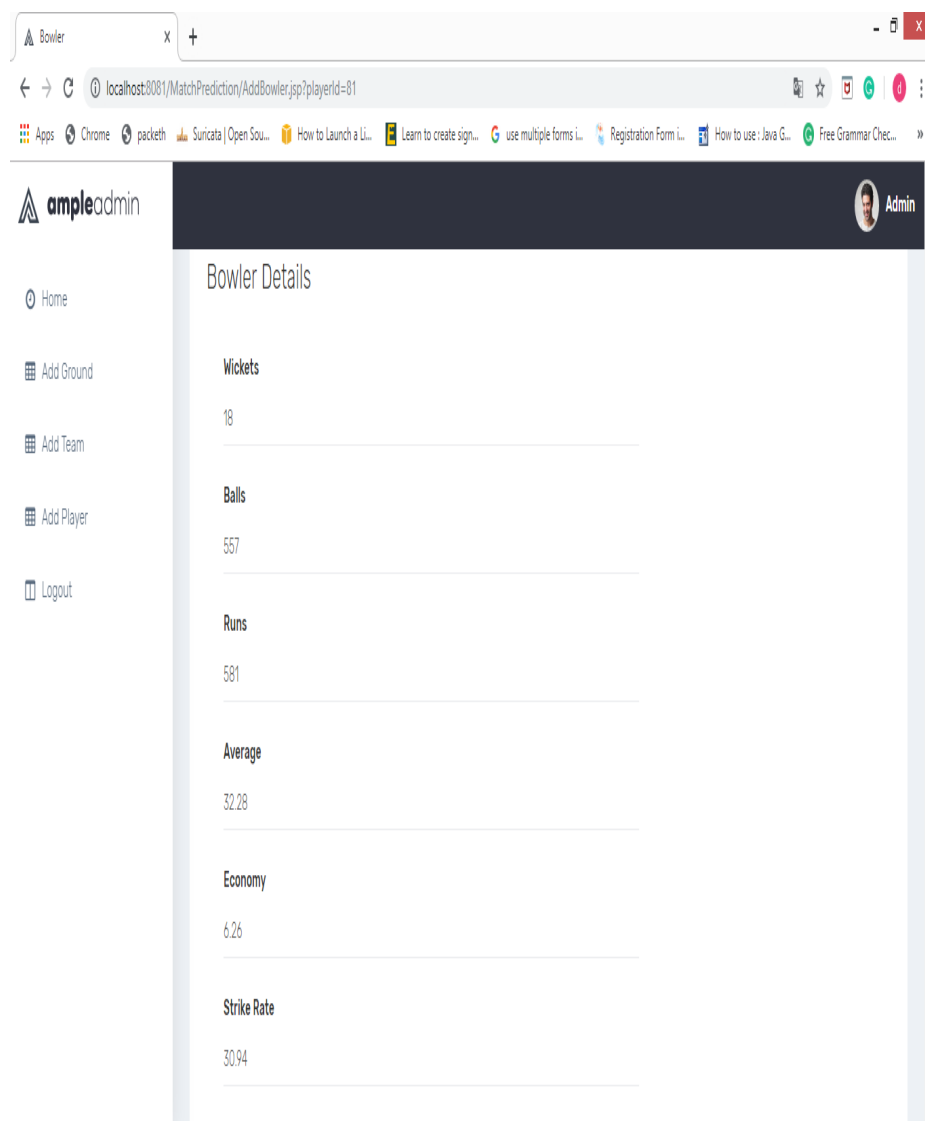


Figure 14: Bowler Details

If the player is All Rounder

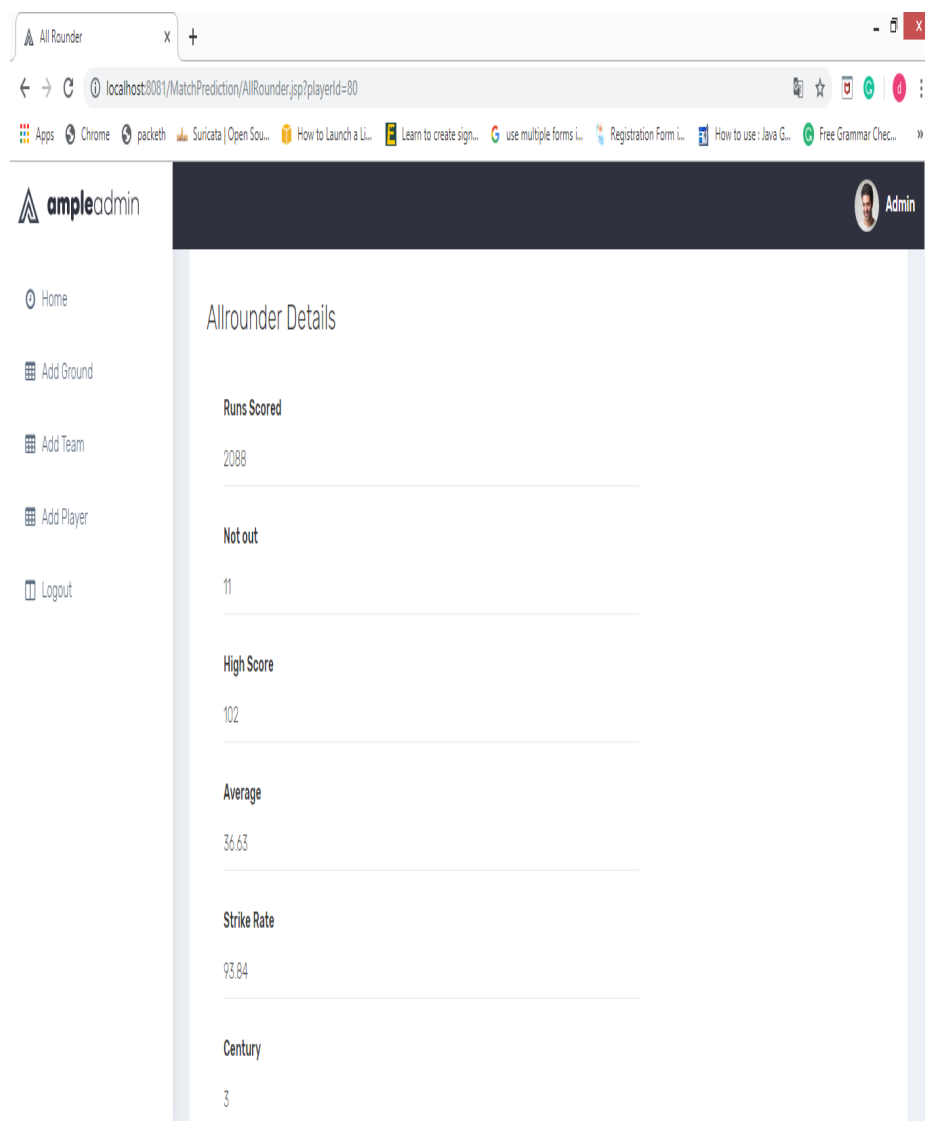


Figure 15: All Rounder(i)

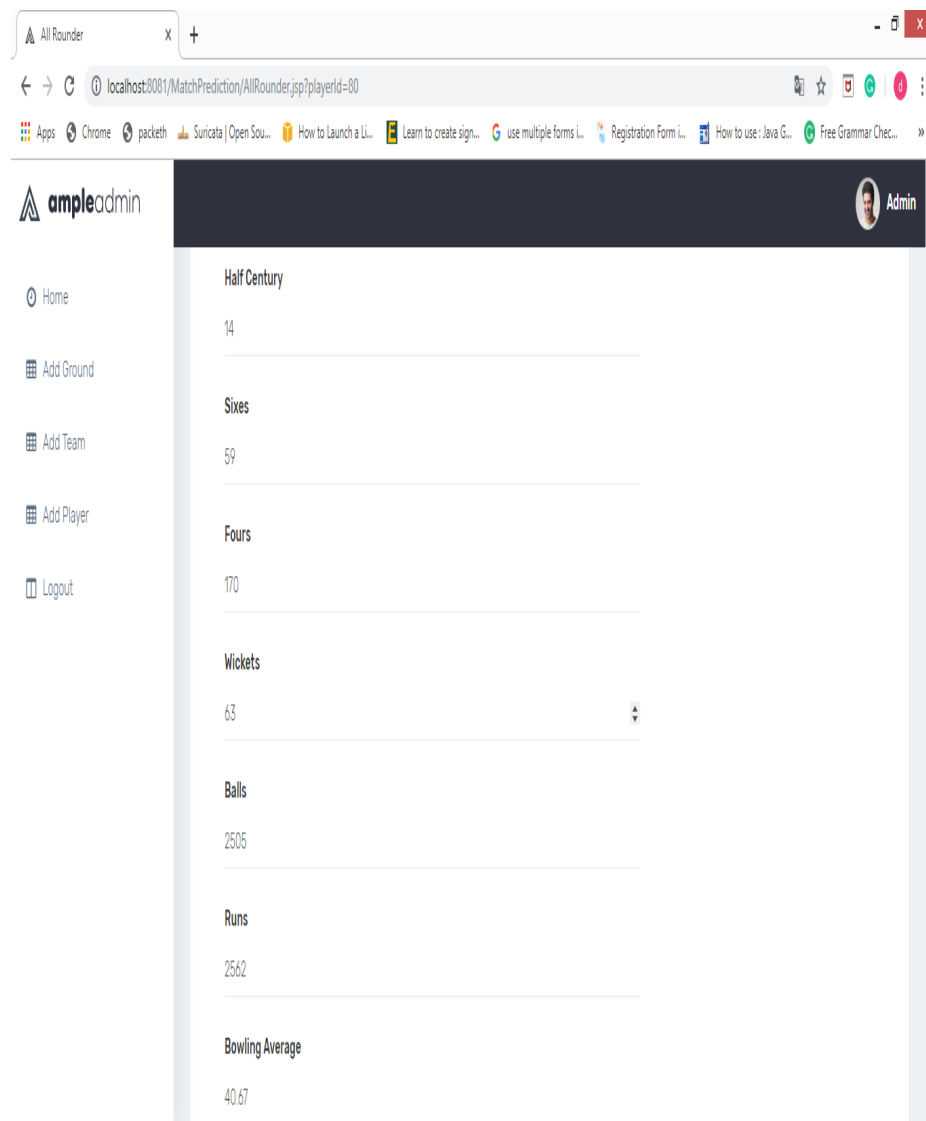


Figure 16: All Rounder(ii)

Add Team

The screenshot shows a web browser window with the address bar displaying 'localhost:8081/MatchPrediction/AddTeam.jsp'. The browser's tab is labeled 'Add Team'. The application interface features a dark blue header with the 'ampleadmin' logo on the left and a user profile labeled 'Admin' on the right. A vertical sidebar on the left contains navigation links: 'Home', 'Add Ground', 'Add Team' (which is highlighted with a blue bar), 'Add Player', and 'Logout'. The main content area is titled 'ADD TEAM' and contains a 'Team Details' section. This section has a label 'Team Name' above a text input field that contains the text 'India'. Below the input field is a green button labeled 'Add'.

Figure 17: Insert Team

Assign Player to team

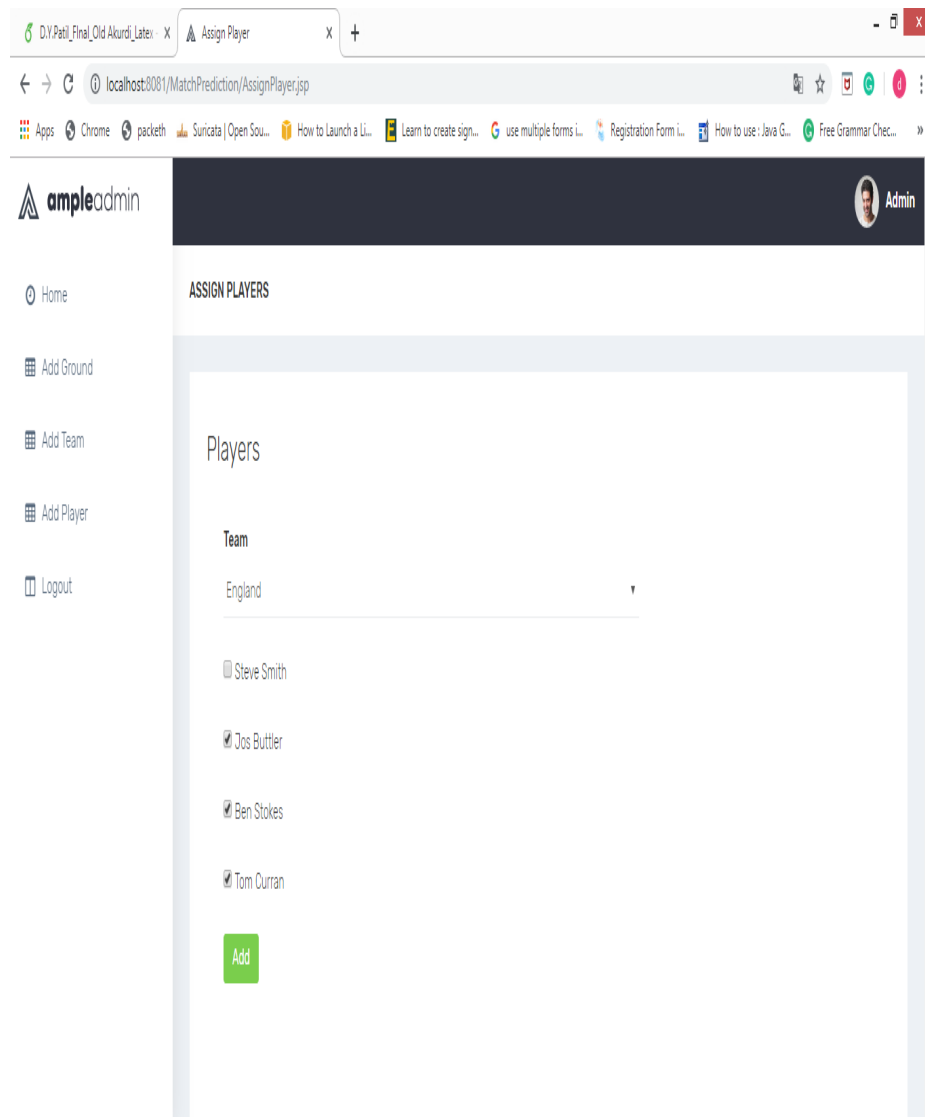


Figure 18: AssignPlayer

Predict winning team UI

The screenshot shows a web browser window with the URL `localhost:8081/MatchPrediction/Predict.jsp`. The browser's address bar and tabs are visible at the top. Below the browser window, the **ampleadmin** dashboard is shown. On the left is a sidebar menu with options: Home, Add Ground, Add Team, Add Player, Assign Player, Playing Eleven, and Logout. The main content area is titled **PREDICT RESULT** and contains a form with the following fields:

- Teams**
 - Team One**: A dropdown menu with `India` selected.
 - Team Two**: A dropdown menu with `New Zealand` selected.
- Ground Name**: A dropdown menu with `Wankhede Stadium` selected.
- Predict**: A green button at the bottom of the form.

Figure 19: Predict UI

Winning team Result

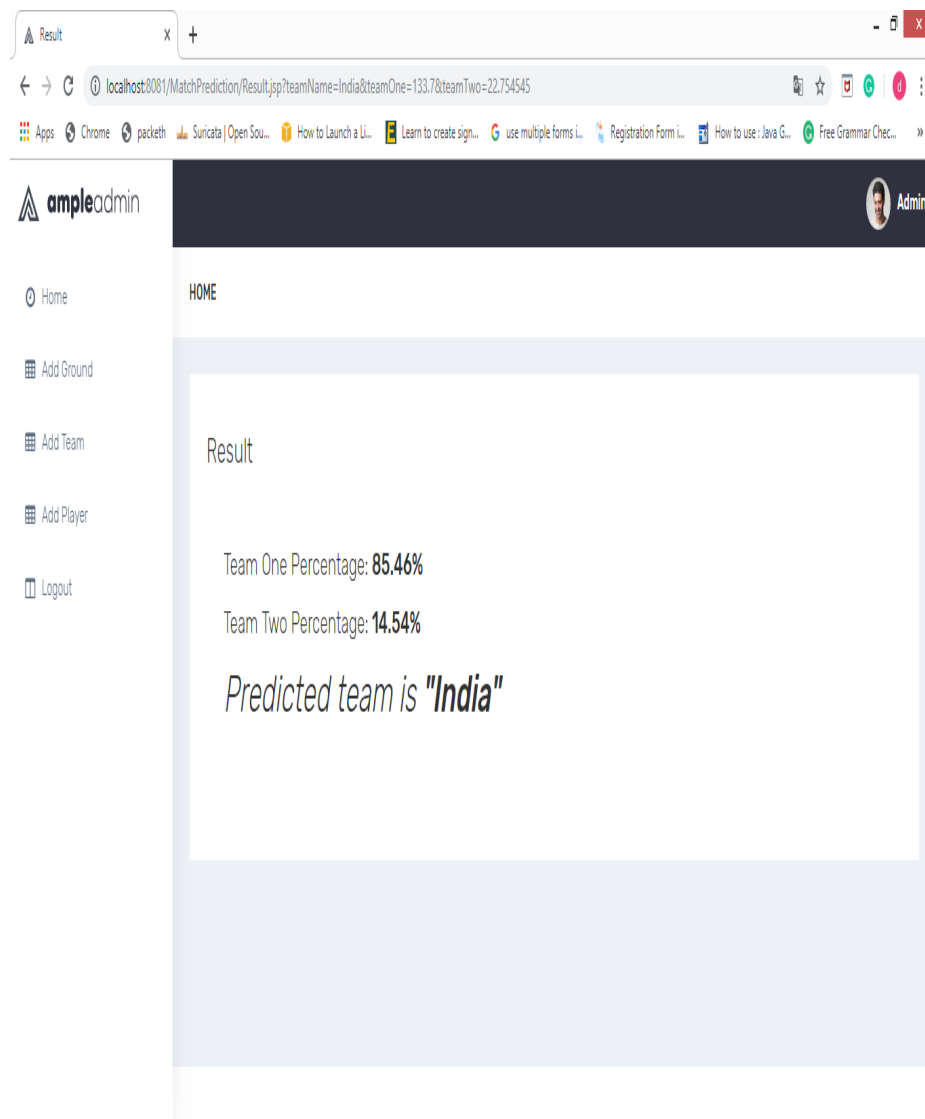


Figure 20: Result

Predict the playing 11 of a team

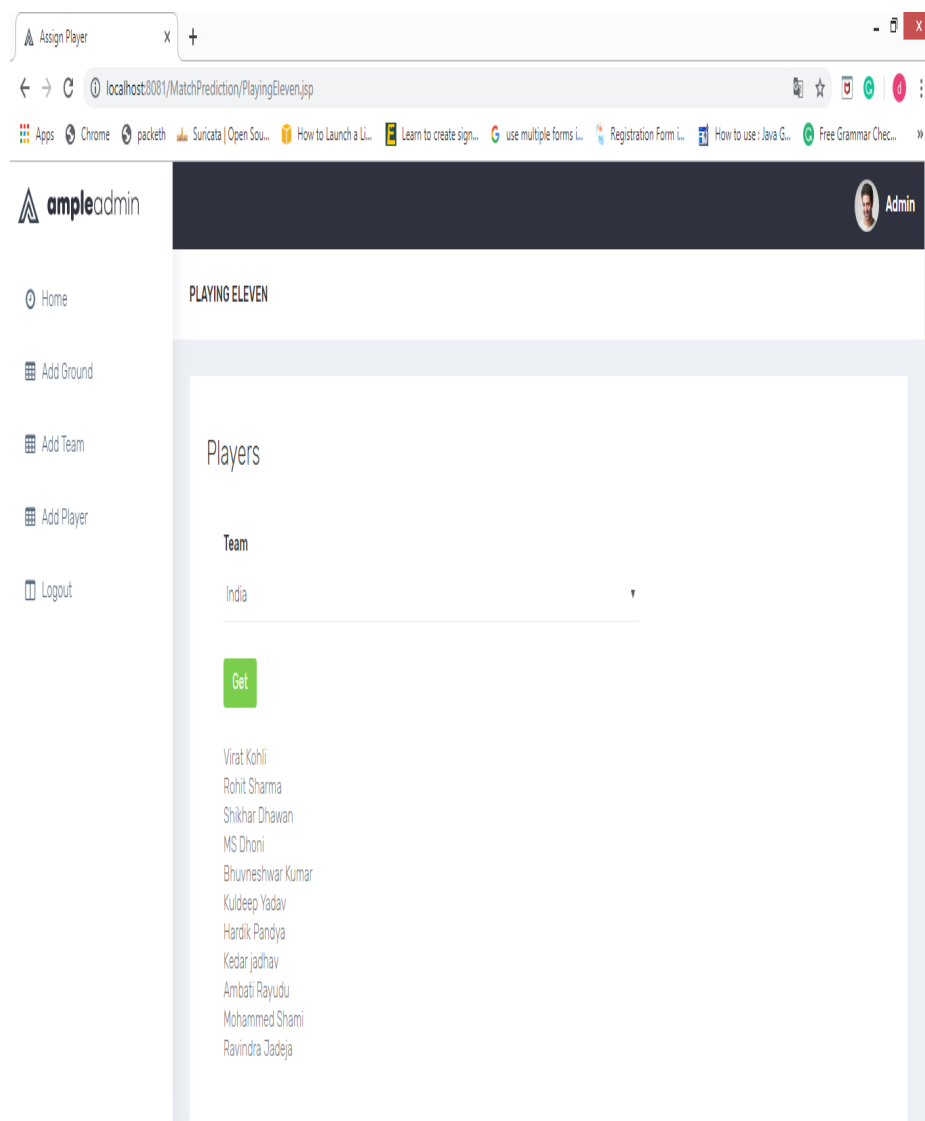
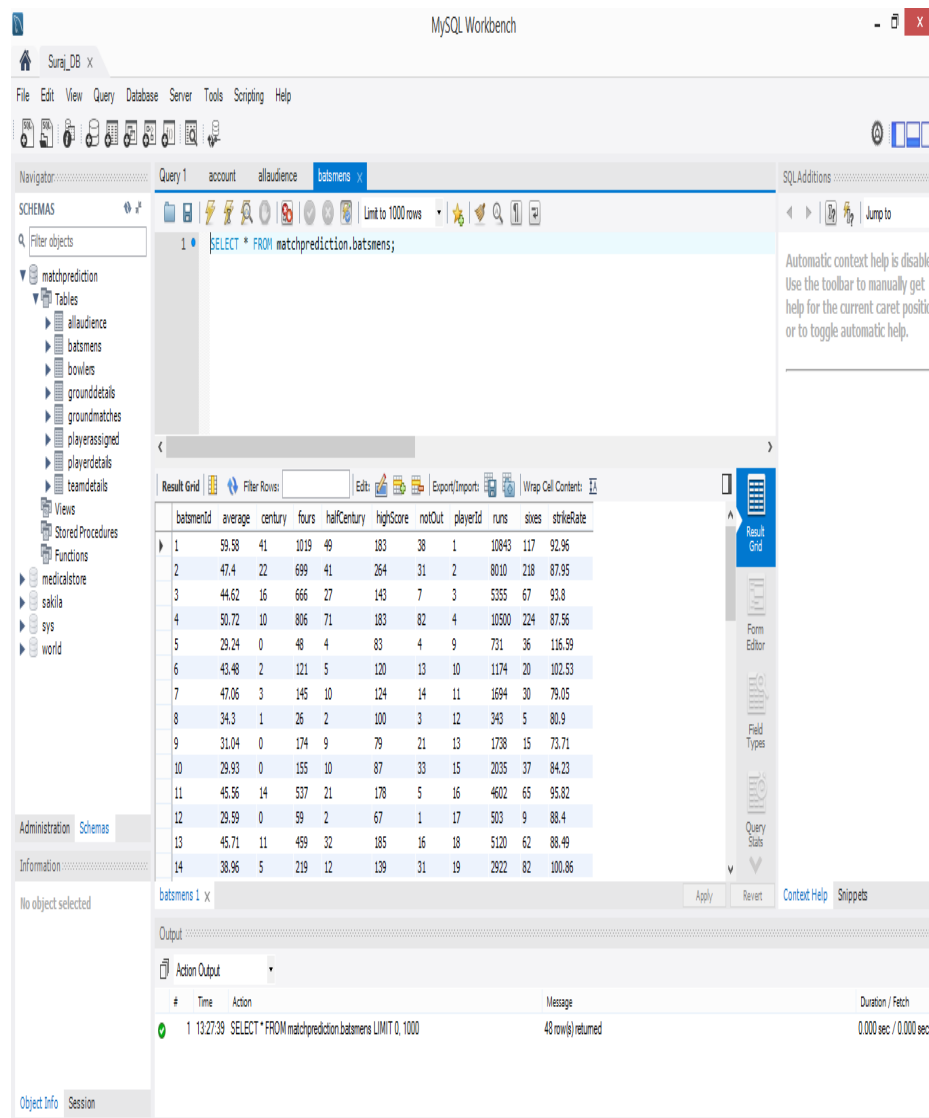


Figure 21: Probable Playing 11

Database



The screenshot shows the MySQL Workbench interface. The 'Schemas' pane on the left lists the database structure, including the 'matchprediction' database with tables like 'allaudience', 'batsmens', 'bowlers', etc. The 'Query Editor' in the center contains the SQL query: `SELECT * FROM matchprediction.batsmens;`. The 'Result Grid' below the query shows 14 rows of data. The 'Output' pane at the bottom shows the execution log with the message: '1 13:27:39 SELECT * FROM matchprediction.batsmens LIMIT 0, 1000 49 row(s) returned 0.000 sec / 0.000 sec'.

batsmenId	average	century	fours	halfCentury	highScore	notOut	playerId	runs	sixes	strikeRate
1	59.58	41	1019	49	183	38	1	10843	117	92.96
2	47.4	22	699	41	264	31	2	8010	218	87.95
3	44.62	16	666	27	143	7	3	5355	67	93.8
4	50.72	10	806	71	183	82	4	10500	224	87.56
5	25.24	0	48	4	83	4	9	731	36	116.59
6	43.48	2	121	5	120	13	10	1174	20	102.53
7	47.06	3	145	10	124	14	11	1694	30	79.05
8	34.3	1	26	2	100	3	12	343	5	80.9
9	31.04	0	174	9	79	21	13	1738	15	73.71
10	29.93	0	155	10	87	33	15	2035	37	84.23
11	45.56	14	537	21	178	5	16	4602	65	95.82
12	29.39	0	59	2	67	1	17	503	9	88.4
13	45.71	11	459	32	185	16	18	5120	62	88.49
14	38.96	5	219	12	139	31	19	2922	82	100.86

Figure 22: Database

8 Deployment and Maintenance

8.1 Operational Requirements to run the System

8.1.1 Installing Software

This system requires Apache Tomcat framework to be installed in the system with the version 7.0. When a user is installing this software on their laptop, it may help to know that this applications will run on JAVA IDE. Once you load the project other packages,jar les, however, may require you to load in the project from storage. Also, when installing software, make sure that the package is labeled forWindows PCs.To install software,open the project in Eclipse and select the Run option.

Installation Steps

1 Copy jdk 1.7.80 to

c: → programfiles(x86) → Java → jdk1.7.80

set the path of bin folder located in jdk in system variables 'path' variable

then goto cmd and enter command "java -version" to check wether jdk setu is done or not

2 Extract eclipse, apache tomcat, mysqlquery browser an install xamp server if mysql is not installed on system

3 Now open eclipse to set up server

goto windows → preferances → server → select runtime enviroment now select apache7.0 and click next

set the path where apache is extracted on your system

set systems jdk path

finish

4 Now select dynamic web project from new project and you can see there runtime enviroment as you setup

mysql db restore and backup.

Installing MySQL:-

- Windows con
gures MySQL Installer.
- A welcome screen provides several options.
- Choose the
rst option: Install MySQL Products.
- Download the latest MySQL products : MySQL installer checks and downloads the latest MySQL products including MySQL server, MySQL Workbench, etc.
- Choosing a Setup Type: there are several setup types available. Choose the Full option to install all MySQL products and features.
- Installation Progress: MySQL Installer downloads all selected products. It will take a while, depending on which products that you selected and the speed of your internet connection.
- MySQL Server Con
guration: choose Windows service details including Windows Service Name and account type, then click Next button to continue.
- Installation Completes: the installation completes. Click
nish button to close the installation wizard and launch the MySQL
Workbench. TSSM's

8.1.2 Uninstalling Software

To uninstall a piece of software from your laptop,first open the Eclipse and select the project which we want to delete and right click on it. Follow the prompts that appear on the screen to finish uninstalling the software.

8.2 User help

- Register and Login to the system.
- When you logged in as an Admin, you will be having access to manage the users in the system.
- Admin is the user who can monitor and control complete system.

9 Conclusion and Future Scope

9.1 Conclusion and Future Scope

This project is undertaken to design a predictive model and evaluate outcome. Using this system we can easily know which team have major chances to win or lose. Prediction system for other games such as Hockey, Baseball and Basketball can also be created. But for that we have to choose different attributes which are significant in their respective game. Such systems simply depend on the selection of right features. In future, this work can be modified on the assumptions made in the starting that every player is fit and their training data is not available, so using training data will increase the accuracy of the prediction. By analyzing different attributes related to the ODI game, we have been able to predict the winning criteria formulated using attributes such as ground match performance, player performance statistics and etc from the dataset.

10 References

1. Amal Kaluarachchi and Aparna S. Varde “CricAI: A Classification Based Tool to Predict the Outcome in ODI Cricket”, Department of Computer Science Montclair State University, Montclair, NJ, USA, vol 15 pp. 250–255, 2010.
2. Tejinder singh, Vishal Singla and Parteek Bhatia, “Score and Winning Prediction in Cricket through Data Mining”, students of Computer Science Engineering, Thapur University Patiyala, Panjab, India (ICSCTI) Oct 8,2015.
3. Bharati S. Kannolli, Prabhu R. Bevinmarad, “Analysis and Prediction of Sentiments for Cricket Tweets Using Hadoop” M. Tech Student, Dept. of Computer Science Engineering, Vijaypur, Karnataka, India (IRJET) Volume: 04 Issue: 10 — Oct -2017.
4. Pranavan Somaskandhan, et al., “Identifying the Optimal Set of Attributes that Impose High Impact on the End Results of a Cricket Match Using Machine Learning”, Department of Computer Engineering, University of Peradeniya, 2017.
5. Shubham Agarwal et al., “Cricket Team Prediction with Hadoop: Statistical Modelling Approach”, Jaypee Institute of Information Technology, Noida, India, 2017.
6. Raza Ul Mustafa et al., “Predicting The Cricket Match Outcome Using Crowd Opinions On Social Networks” Department of Computer Science, COMSATS Institute of Information Technology, 57000 Sahiwal Campus, Pakistan, 2017.
7. <http://ieeexplore.ieee.org/Xplore/home.jsp>
8. <https://en.wikipedia.org/wiki/NaiveBayesclassifier>
9. <https://en.wikipedia.org/wiki/Machinelearning>
10. <https://towardsdatascience.com/sentiment-analysis-with-pyspark-bc8e83f80c35>
11. <https://dzone.com/articles/ipl-cricket-analytics-and-predictive-model>

Appendix A Testing Of Project Problem Statement Using Generated Test Data:

- Testing of project problem statement using generated test data (using mathematical models, GUI, Function testing principals) selection and appropriate use of testing tools, testing of UML diagrams reliability.
- **Testing:** We intend to design and check proper connection of backend and frontend, division of dataset into historical and training data. Proper display of all the graphs.
- **Requirement Testing:** A common misconception of software testing is that its only conducted after a system has been completed. However, software testing should be a continuous process that takes place throughout the development life cycle of a system. Hence we began testing our proposed system design to check if it meets its desired functionality.
- **Unit testing:** The objective of unit testing is to isolate a section of code and verify its correctness. In this the smallest testable parts of an application called units, are individually and independently tested.

The proposed system intends to present results essential to a full understanding of system performance. The following parameters are considered:

1. GUI response
2. Response Time

- **Formal Testing of GUI**

1. **Testing of User Login GUI:**

- The user should be able to create his own personal account based upon his unique loginID and Password.

2. **Testing of Data Insertion:**

- The user must be able to enter data into his own specific account.
- Data of types including Text, and Numbers should be entered with ease.

3. Validation of user's data and graphs based upon the analytics:

- To provide correct results of analytics on various grounds.

Appendix B Project Planner

NO	TASK	DURATION (Days)	START DATE	END DATE
1	Group Formation	4	21 Jun	26 Jun
2	Decide Area Of Interest	4	26 Jun	28 Jun
3	Search Topic	5	28 Jun	6 Jul
4	Topic Selection	5	6 Jul	16 Jul
5	Topic Selection	5	16 Jul	22 Jul
6	Search Related Information	12	22 Jul	12 Aug
7	Understanding Concept	7	12 Aug	20 Aug
8	Search Essential Document(IEEE and White Paper,) Software)	6	20 Aug	31 Aug
9	Problem Definition	2	31 Aug	8 Sep
10	Literature Survey	5	8 Sep	18 Sep
11	SRS	14	18 Sep	20 Sep
12	Project Planning	2	20 Sep	30 Sep
13	Modeling and design	10	30 Sep	2 Oct
14	Technical Specification	2	30 Sep	2 Oct
15	PPT	6	2 Oct	5 Oct