

CAPSTONE PROJECT REPORT

(Project Term January-May 2021)

PREDICTIVE ANALYSIS ON IPL: INDIAN PREMIER LEAGUE

Submitted by

Gulladurti Tarunkanth Reddy
Beeravalli Vamshi Reddy
Routhu Praveen
Majji Naveen Kumar

Registration Number: 11705270
Registration Number: 11705309
Registration Number: 11701913
Registration Number: 11703801

Project Group Number: CSERGC0225

Course Code: CSE445

Under the Guidance of

Dr. Baljit Singh Saini

School of Computer Science and Engineering



L OVELY
P ROFESSIONAL
U NIVERSITY

TOPIC APPROVAL PERFORMANCE

School of Computer Science and Engineering (SCSE)

Program : P132::B.Tech. (Computer Science & Engineering)

COURSE CODE : CSE445

REGULAR/BACKLOG : Regular

GROUP NUMBER : CSERGC0225

Supervisor Name : Dr. Baljit Singh Saini

UID : 22078

Designation : Associate Professor

Qualification : _____

Research Experience : _____

SR.NO.	NAME OF STUDENT	Prov. Regd. No.	BATCH	SECTION	CONTACT NUMBER
1	Beeravalli Vamshi Reddy	11705309	2017	K17DQ	7780708071
2	Gulladurti Tarunkanth Reddy	11705270	2017	K17DQ	9963663814
3	Majji Naveen Kumar	11703801	2017	K17DW	8374059752
4	Routhu Praveen	11701913	2017	K17DW	9347036002

SPECIALIZATION AREA : System Programming

Supervisor Signature: _____

PROPOSED TOPIC : Predictive Analysis on IPL- Indian Premier League

Qualitative Assessment of Proposed Topic by PAC		
Sr.No.	Parameter	Rating (out of 10)
1	Project Novelty: Potential of the project to create new knowledge	7.80
2	Project Feasibility: Project can be timely carried out in-house with low-cost and available resources in the University by the students.	8.40
3	Project Academic Inputs: Project topic is relevant and makes extensive use of academic inputs in UG program and serves as a culminating effort for core study area of the degree program.	8.00
4	Project Supervision: Project supervisor's is technically competent to guide students, resolve any issues, and impart necessary skills.	8.80
5	Social Applicability: Project work intends to solve a practical problem.	7.80
6	Future Scope: Project has potential to become basis of future research work, publication or patent.	7.80

PAC Committee Members		
PAC Member (HOD/Chairperson) Name: Pushpendra Kumar Pateriya	UID: 14623	Recommended (Y/N): Yes
PAC Member (Allied) Name: Varun Dogra	UID: 12343	Recommended (Y/N): Yes
PAC Member 3 Name: Dr. Baljit Singh Saini	UID: 22078	Recommended (Y/N): Yes

Final Topic Approved by PAC: Predictive Analysis on IPL- Indian Premier League

Overall Remarks: Approved

PAC CHAIRPERSON Name: 13897::Dr. Deepak Prashar

Approval Date: 16 Mar 2021

DECLARATION

We hereby declare that the project work entitled “Predictive Analysis on IPL – Indian Premier League” is an authentic record of our own work carried out as requirements of Capstone Project for the award of B. Tech degree in Computer Science and Engineering from Lovely Professional University, Phagwara, under the guidance of Dr. Baljit Singh Saini, during January to May 2021. All the information furnished in this capstone project report is based on our own intensive work and is genuine.

Project Group Number: CSERGC0225

Name of Student 1: Gulladurti Tarunkanth Reddy

Registration Number: 11705270

Name of Student 2: Beeravalli Vamshi

Registration Number: 11705309

Name of Student 3: Routhu Praveen

Registration Number: 11701913

Name of Student 4: Majji Naveen Kumar

Registration Number: 11703801

(Signature of Student 1)

Date:

(Signature of Student 2)

Date:

(Signature of Student 3)

Date:

(Signature of Student 4)

Date:

CERTIFICATE

This is to certify that the declaration statement made by this group of students is correct to the best of my knowledge and belief. They have completed this Capstone Project under my guidance and supervision. The present work is the result of their original investigation, effort and study. No part of the work has ever been submitted for any other degree at any University. The Capstone Project is fit for the submission and partial fulfillment of the conditions for the award of B. Tech degree in Computer Science and Engineering from Lovely Professional University, Phagwara.

Signature and Name of the Mentor

Designation

School of Computer Science and Engineering,
Lovely Professional University,
Phagwara, Punjab.

Date :

ACKNOWLEDGEMENT

We are using this opportunity to express our gratitude to everyone who supported us throughout this capstone project. This project would be incomplete without the mentioning of the people whose constant guidance and encouragement made it possible. We take pleasure in presenting before you, our project, which is a result of the studied blend of both research and knowledge. We are sincerely grateful to them for sharing their truthful and illuminating views on several issues related to the project.

We express our earnest gratitude to our mentor Dr. Baljit Singh Saini for his constant support, encouragement and guidance. We are grateful for his cooperation and his valuable suggestions.

We express our thanks to the Head of the Department and College Management for all their support and encouragement. Finally, we express our gratitude to all other members who involved either directly or indirectly for the completion of this project.

Group Members

Gulladurti Tarunkanth Reddy

Beeravalli Vamshi Reddy

Routhu Praveen

Majji Naveen Kumar

TABLE OF CONTENTS

Inner first page.....	(i)
PAC form.....	(ii)
Declaration.....	(iii)
Certificate.....	(iv)
Acknowledgement.....	(v)
Table of Contents.....	(vi)
1. Introduction	1
1.1 Overview	1
1.2 History	2
2. Scope of the Study	3
3. Existing System	3
3.1 Introduction	3
3.2 Existing Software	4
3.3 DFD for present system	8
3.4 What's new in the system to be developed	10
4. Problem Analysis	11
4.1 Product definition	11
4.2 Feasibility Analysis	11
4.3 Project Plan	12
5. Design	12
5.1 System Design	12
5.2 Detailed Design	13
5.3 Flowcharts	15
5.4 Pseudo code	17
6. Testing	23
6.1 Functional Testing	23
6.2 Structural Testing	23
6.3 Levels of Testing	24
6.4 Verification and Validation	25

7. Project Results	26
7.1 Data Analysis	26
7.2 Dash Boards	28
7.3 Algorithms	29
7.4 Predictions	36
7.5 Web Interface	37
8. Implementation	39
9. Conclusion	42
10. References	42

1. Introduction

1.1 Overview

Cricket is a bat-and-ball sport in which two teams of eleven players compete on a field with a 22-yard (20-meter) pitch in the middle and two wickets at each end, each with two bails balanced on three stumps. The batting side scores runs by hitting the ball bowled at the wicket with the bat (and running between the wickets), whereas the bowling and fielding teams attempt to prevent this (by preventing the ball from leaving the field and bringing it to either wicket) and dismissal of each batsmen (so they are "out"). Being bowled, when the ball strikes the stumps and dislodges the bails, and by the fielding side catching the ball after it has been struck by the bat and before it hits the pitch, or hitting a wicket with the ball before a batter can reach the crease in front of the wicket, are both methods of dismissal. The innings ends when ten batters have been removed, and the teams switch positions. In international matches, a third umpire and match referee assist two umpires in evaluating the game.



Figure 1.1: Batting Pitch

Cricket matches range from Twenty20, in which each team bats for a single 20-over innings, to Test matches, which plays for last five days. Cricketers traditionally wear all-white uniforms, but in limited overs cricket, they wear club or squad colours. Some players wear protective gear in addition to their basic gear to avoid injury from the ball, which is a hard, strong spheroid of a cork core which is encased in compressed leather with a slightly raised sewn seam and layered with tightly wound string.

The Indian Premier League (IPL) is a best and professional Twenty20 cricket league in India that has began since 2008. Teams from major Indian cities compete in the league, which is based on a round-robin group and knockout format. The IPL, which was created by the Board of Control for Cricket in India (BCCI), has grown to be the most lucrative and successful cricket league in the world. Matches usually start in the late afternoon or evening, with at least a portion of them being played under floodlights at night to increase the television audience for international broadcasts. The league used to be played on a home and away basis between all teams, but with the planned expansion to ten clubs (divided into two groups of five) in 2011, the format was modified so that matches between certain teams would be limited to a single match. The top four teams compete in three play-off matches, with one losing team getting a second chance to advance to the final, a twist designed to increase future television revenue. The four teams that finished at the top of the tables compete in a series of elimination games in the play-off portion of the competition, which gives one team that lost its first-round game a second chance to progress to the final match.

1.2 History

Zee Entertainment Enterprises funded the creation of the Indian Cricket League (ICL) in 2007. The Board of Control for Cricket in India (BCCI) and the International Cricket Council (ICC) did not recognise the ICL, and the BCCI was unhappy with its committee members joining the ICL executive board. To discourage players from entering the ICL, the BCCI increased prize money in domestic tournaments and enforced lifetime bans on players who joined the ICL, which the board considered as a rebel league.

In the IPL, the world's best cricketers became millionaires almost overnight, despite seldom earning as much as their counterparts in other professional sports. Major companies, Bollywood stars, and media moguls were among the IPL franchise owners who bid for the best players in the league's auctions. The well-funded Mumbai Indians had the league's largest payroll, more than \$100 million, at the start of the IPL. In the initial auction for the 2008 season, the Chennai Super Kings paid \$1.5 million for

Mahendra Dhoni's services, and the Kolkata Knight Riders paid \$2.4 million for Gautam Gambhir, the Indian national team's opening batsman, in the bidding for the 2011 season.

The Mumbai Indians (MI), Chennai Super Kings (CSK), Royal Challengers Bangalore (RCB), Deccan Chargers (DC which is based in Hyderabad), Delhi Daredevils (DD), Punjab XI Kings ((PXXIK) Mohali), Kolkata Knight Riders (KKR), and Rajasthan Royals ((RR) Jaipur) were the eight founding franchises. The BCCI removed two franchises, Rajasthan and Punjab, from the league in late 2010 due to ownership policy violations, but they were later reinstated in time for the 2011 tournament. The IPL welcomed two new franchises for the 2011 season: the Pune Warriors India and the Kochi Tuskers Kerala. The BCCI terminated the Kochi club's contract after just one year of service. The Deccan Chargers were replaced by the Sunrisers Hyderabad in the IPL in 2013.

2. Scope of the Study

The study is about to start a new way of generation where each and every cricket fan should approach this method for all the tournaments including country wise and state wise like our project and the main important point is to give a full detailed view in IPL tournament including analysis and predictions in an interface. Also, the illegal activities like gambling, betting on score and win, ...etc are now reformed as legal activities like dream11, bet360, etc. With our project all these activities are made easy.

3. Existing System

3.1 Introduction

Python is a good language for prototyping because it allows you to switch out modules easily. Python is one of the most user-friendly programming languages available, thanks to its simple syntax and removal of complication, which places a greater focus on natural language. Python codes can be written and implemented much faster than other programming languages due to their ease of learning and use. Python is one of the most open programming languages available because it has a simple syntax and is not overly complex, allowing natural language to grow stronger.

Python codes can be written and executed much faster than other programming languages due to their ease of learning and use.

A feature is an individual measurable property or attribute of a phenomenon being observed in machine learning and pattern recognition. Pattern recognition, classification, and regression algorithms include the selection of descriptive, discriminating, and independent features.

The transportation industry relies on making routes more effective and anticipating future issues to improve profitability, so analysing data to detect patterns and trends is critical. Machine learning's data processing and modelling capabilities are valuable resources for distribution firms, public transport, and other transportation organizations.

3.2 Existing Software

3.2.1 Some Pros of Python

1. Readable and Secure Code

When writing a software application, you should focus on the quality of its source code to facilitate maintenance and updating. Python syntax rules allow you to express concepts without writing additional code. In contrast to other programming languages, Python prioritizes code readability and allows you to use English keywords instead of punctuation marks. Therefore, you can use Python to build applications without writing additional code. A readable and clean code base will help you maintain and update the software without spending extra time and effort.

2. Multiple Programming Paradigms

Python, like other modern programming languages, has a variety of programming applications which supports. It supports fully structured and fully structured production. Also, its language features support a variety of concepts in practical and objective programs. At the same time, Python also has a powerful system and

automatic memory management. Editing paradigms and language features help you use Python to build large and complex software applications.

3. Compatible with major platforms and programs

Currently, Python supports many running applications. You can even use Python translators to apply the code to certain platforms and tools. Also, Python is a translated programming language. Allows you to apply the same code to many platforms without refund. Therefore, you do not need to re-update the code after making the change. You can activate the modified code of the application without refund and check the effect of changes made to the code immediately. The feature makes it easy for you to make changes to the code without increasing the development time.

4. General Library

Its large and sturdy library makes Python score more points than any other programming language. A standard library allows you to choose from multiple modules depending on your specific needs. Each module allows you to add functionality to a Python application without having to write additional code. For example, when writing a web application in Python, you can use certain modules to run web applications, perform cable tasks, manage the operating interface or work through online contracts. You can collect information on various modules by browsing Python Standard Library documents.

5. Multiple Frameworks and Open-Source Tools

As an open-source programming language, Python helps you reduce the cost of software development significantly. You can use several open-source Python frameworks, libraries and development tools to reduce development time without increasing development costs. You have the option to choose from a wide range of Python frames and development tools depending on your specific needs. For example, you can simplify and speed up web application development by using Python web frameworks such as Django, Flask, Pyramid,

Bottle and CherryPy. Similarly, you can speed up desktop application development using Python GUI frameworks and toolkit tools such as PyQt, PyJs, PyGUI, Kivy, PyGTK and WxPython.

6. Simplify complex software development

Python is a general programming language. Therefore, you can use programming language to improve desktop and web applications. Also, you can use Python to make complex science programs and numbers. Python is built with features that facilitate data analysis and visualization. You can take advantage of Python's data analysis features to create great custom data solutions without spending extra time and effort. At the same time, Python's data-science libraries and APIs help you visualize and present information in a very attractive and effective way. Many Python developers use Python to implement artificial intelligence (AI) and natural language correction functions.

7. Embrace Test Process Development

You can use Python to create a type of software application quickly. Also, you can create a software application directly from the prototype by simply using the Python code. Python makes it easy for you to create codes and experiments simultaneously using the advanced test method (TDD). You can easily write the required tests before writing the code and then use the tests to check the program code continuously. The test can be used to check if the app meets the requirements specified based on its source code.

3.2.2 Django

Django may be the last post-development framework for website design and construction, and could be the basis for the increase in quality Python's over the years. Django uses a model-view-template (MVT) design, the pattern supports a group of best practices for editing your code.

Model: Model is what binds your app to information. usually, the model represents a table within the information and describes how your application accesses the information. Simplifies tasks to create, delete, or change table entries.

View: Lessons for computer programming. It provides templates and describes what you see and therefore its functionality. built with HTML, CSS, and JavaScript files.

Template: that static HTML files you want with a special syntax. Defines embedded content.

3.2.3 HTML, CSS

HTML (Hypertext Markup Language) and CSS (Cascading Style Sheets) are two of the most popular Web page building technologies. For a number of devices, HTML provides the page structure and CSS provides the (visual and aural) layout. HTML and CSS, along with graphics and scripting, are the foundations for creating Web pages and Web applications.

The language used to describe the layout of Web pages is HTML. HTML gives authors the means to:

- Create online documents that include headings, text, tables, lists, images, and more.
- At the click of a button, you can access online information via hypertext links.
- Create forms for searching for details, making reservations, ordering items, and other transactions with remote services, etc.
- Spreadsheets, video clips, sound clips, and other applications can all be used in their documents directly.

Authors use markup to explain the layout of pages in HTML. The language's element designate content types such as "paragraph," "list," "table," and so on.

CSS is a language for explaining how Web pages are presented, including colours, layout, and fonts. It enables the presentation to be adjusted for various devices such as computer, such as large screens, small screens, or printers. CSS can be used for any XML-based markup language and is not dependent on HTML. The separation of HTML and CSS makes it easier to manage websites, share style sheets across pages, and modify pages for different situations. The separation of structure (or: content) from presentation is referred to as this.

3.2.4 Machine Learning

In machine learning, supervised learning algorithms are trained on labelled instances, such as an input to get the output. A piece of equipment, may have data points labelled “F” (failed) or “R” (passed runs) which is considered as an example. The learning algorithm is given a set of inputs and the correct outputs, and it learns by comparing its actual output to the correct outputs in order to detect errors. It then makes the required changes to the model. Supervised learning uses patterns to estimate the values of the mark on additional unlabeled data using techniques such as classification, regression, estimation, and gradient boosting. In applications where historical data predicts possible future events, where supervised learning is widely used. It can predict when credit card purchases are likely to be fraudulent, or which insurance customers are likely to file a lawsuit, for example.

3.3 DFD for present system

3.3.1 Level 0 DFD

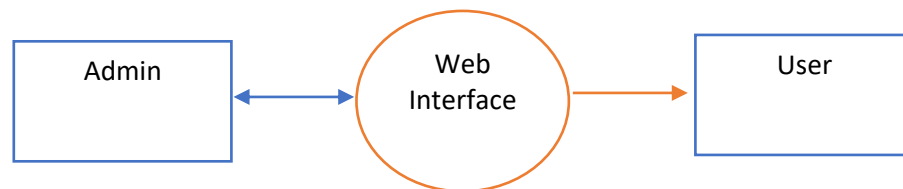


Figure 3.3.1: Level 0 DFD of system

3.3.2 Level 1 DFD

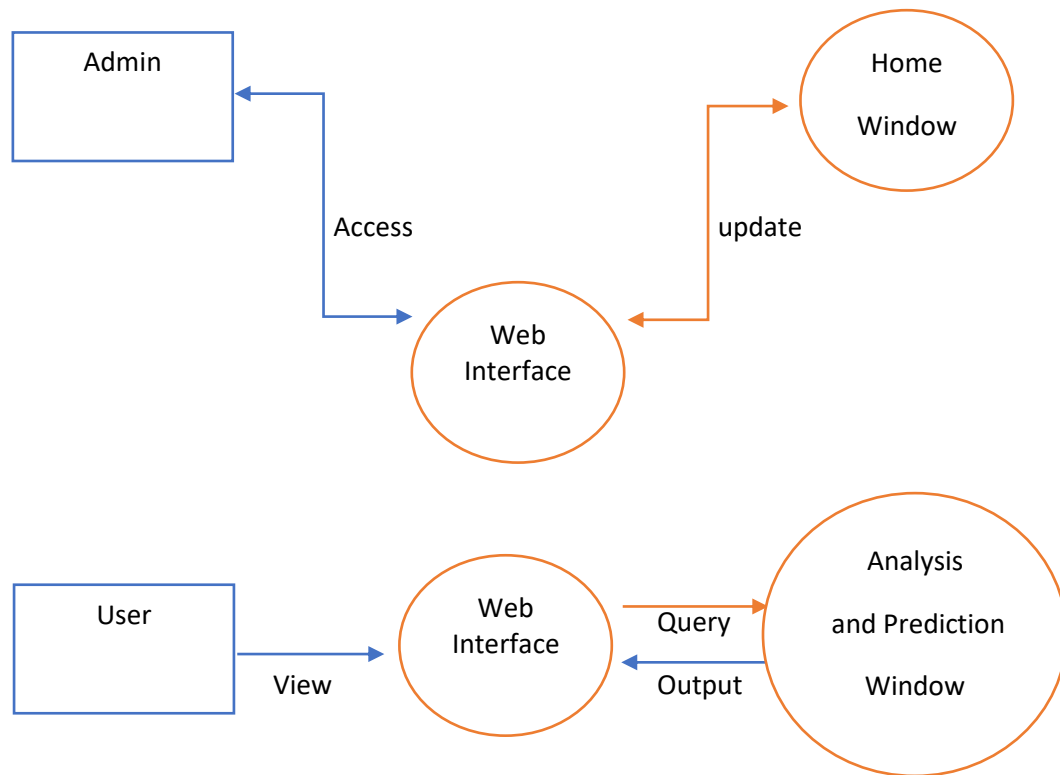


Figure 3.3.2: Level 1 DFD of system

3.3.3 Level 2 DFD

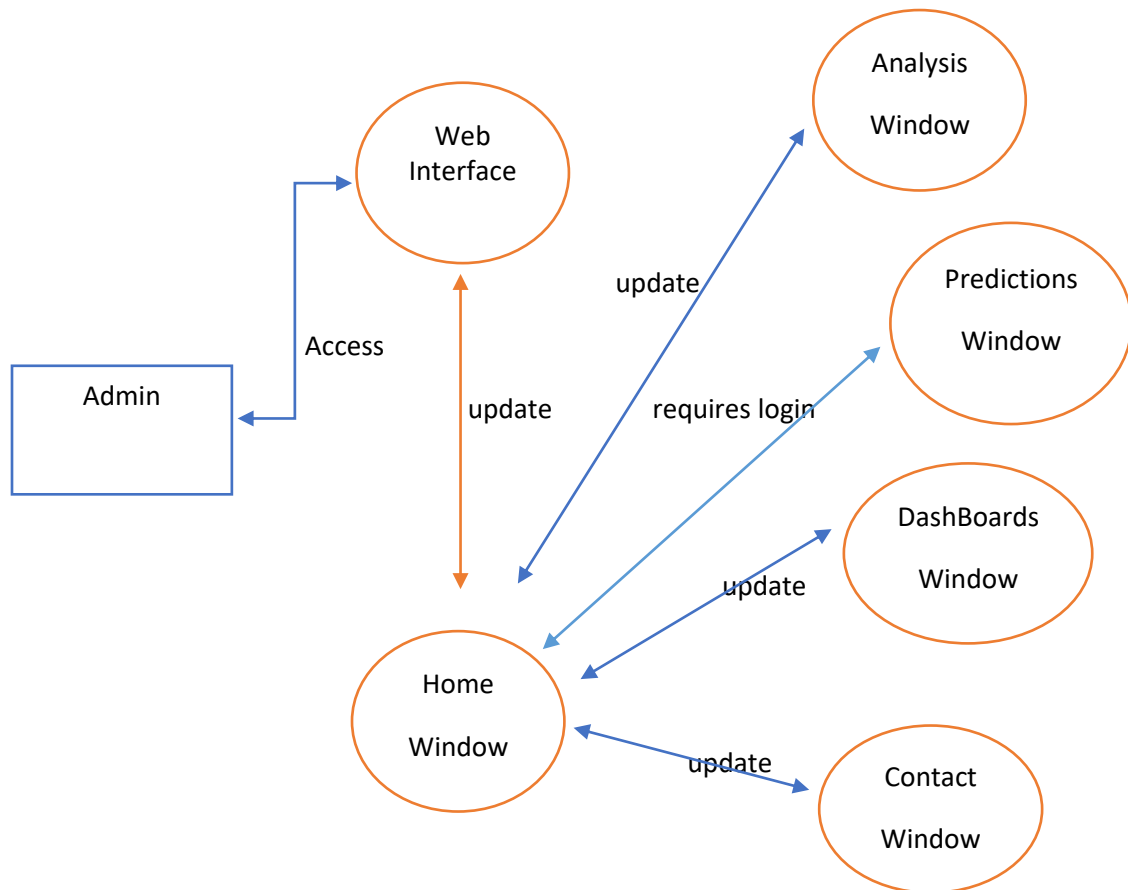


Figure 2.3: Level 2 DFD of system

3.4 What's new in the system to be developed

In this project, there will be a html-based website describing with IPL parameters, which is useful for immediate results for all users. The newness of the project is a HTML based website creation where all the fields are included within the website like Data Analysis, Predictions, Dashboards and importantly the user requires to login if he needs to check the predictions and the user can see the analysis part without login. Also, users are controlled by admins with the help of Django.

In this, there is a particular parameter named Score Prediction which is not similar to other projects of IPL. So, we have decided to include that parameter in our new project.

4. Problem Analysis

4.1 Product definition

Predicting the IPL match is very difficult and predicting the toss, score, player of the match and which team is winning the match and everything is considered for a match. People use different fantasy league apps like dream11, to earn the money by predicting the players performance and winning team without the knowledge about the performance of previous IPL seasons. So, we are providing the smart solution called the IPL predictor which helps you to give the accurate results for the match.

4.2 Feasibility Analysis

The proposed system will not generate precise results every time it is very difficult to predict the IPL match. There are a lot of things that affect the match we cannot predict the match if there is a natural calamity situation happen because it leads to the result differently for example duck worth lousie method which is used for any rainy situation. so they are intended to be unpredictable.

We are taking many possibilities for predicting the match we collected data for the analysis which is used to analyze the players stats, and which team is going to win the match, predicting the team runs, and further it can be continued for predicting the team which is going to win the toss.

The objective of the system is to provide a rough idea about how the IPL works and how it predicts the matches. There are many reasons for not being able to acknowledge the long-term output of a predicted model. Various factors and parameters may influence the prediction due to which prediction is just not feasible.

4.3 Project Plan

4.3.1 Functional Requirements

Functional requirements are the desired operations and features that must be included in any system to satisfy the project requirements that can be acceptable by the users. The functional requirements for the proposed system must require are as follows:

1. The system should provide approximate prediction.
2. The system should collect accurate data from the dataset in a consistent manner.

4.3.2 Non-Functional Requirements

Non-functional requirements are the characteristics and quality attributes of the system that are used to judge the system operations. The non-functional requirements are essentially based on the Data Integrity, Usability, Reliability, security, security and recoverability of the system. The non-functional requirements are as follows:

1. The system should have user friendly interface for an easy usage.
2. It must efficiently at any point of time.
3. The system should be reliable.
4. The system should provide better accuracy.

5. Design

5.1 System Design

The System have been designed used different frameworks of python, HTML, CSS and Tableau public website. The system consists of different pages which are named as:

- Home
- Data Analysis Pages
- Dash Board Pages
- Contact Us
- Feedback Page
- Login Page
- Register Page

These pages are developed using HTML, CSS and Tableau Public website.

- Prediction Pages

The prediction pages are also developed using HTML, CSS where the ML models Pickle data has been connected to the page using Django.

5.2 Detailed Design

The design is explained as how it works when a user controls the interface,

User is able to register at free of charge and logs in, if user logs in then he displays a message as “Welcome username” which is created by user in the process of registration. All these details are stored in Django which can be modified in IP address/admin by admin only. Here IP address is generated by Django python code with the help of command prompt. Also, the user can access the Data Analysis and Dash boards with the help of an internet connectivity where we can filter the analysis same as in tableau desktop. In prediction part the user selects particular parameters which encoded with numericals and in the developed models, it dumps a pickle file which consists of binary numericals where the output is taken as 0 or 1 and displays the decoded message.

Benefits of Using Python in Web Development

There is a square that measures the many factors that transform Python hiring with full development:

- **Easy to learn:** The simplicity of syntax allows you to touch knotty techniques and ensure communication between developers working on the same project is an additional cost-saving option. In fact, the language that is easy to find means that neophyte developers or World Health Organization developers who have no knowledge of python will learn this language and become part of the event team soon.
- **Good readability:** The very fact that python is similar to our everyday language and its pressure on learning means that its engineers will write comprehensible code.
- **Complex tasks at the end:** Python is a complex writing language that allows you to try advanced tasks at the end, as well as AI and Science knowledge on a continuous basis | another} common function in other programming languages.
- **High quality in the middle of complete development:** Sometimes, this feature can be seen as an Associate in Nursing indexing in the appropriate / excellent language, however, this can be a collective bonus for correcting the error of our program. If you are struggling with a specific bug or do not know exactly how to use the replacement feature, someone on the World Health Organization website may already have a similar problem, and getting a response is just a click away.
- **Libraries vary:** There are a number of python library libraries that you can use to speed up your app event. These pre-coded square code libraries are there for everyone, so you don't have to be forced to renew the wheel. a certain number of libraries measure Numpy, scitLearn for data analysis and mathematical algorithms, and x for SQL query integrated.
- **Good Framework:** many of the leading fashion trends for python net development including Django, Flask, Pyramid, Web2Py, and Turbogears. These frames measure the square as tool boxes that will help you speed up an

online program event. They contain standard packages and modules to help you with most application areas (URL routing, accessing the info, HTTP requests, and responses).

5.3 Flowcharts

5.3.1 User Case Diagram

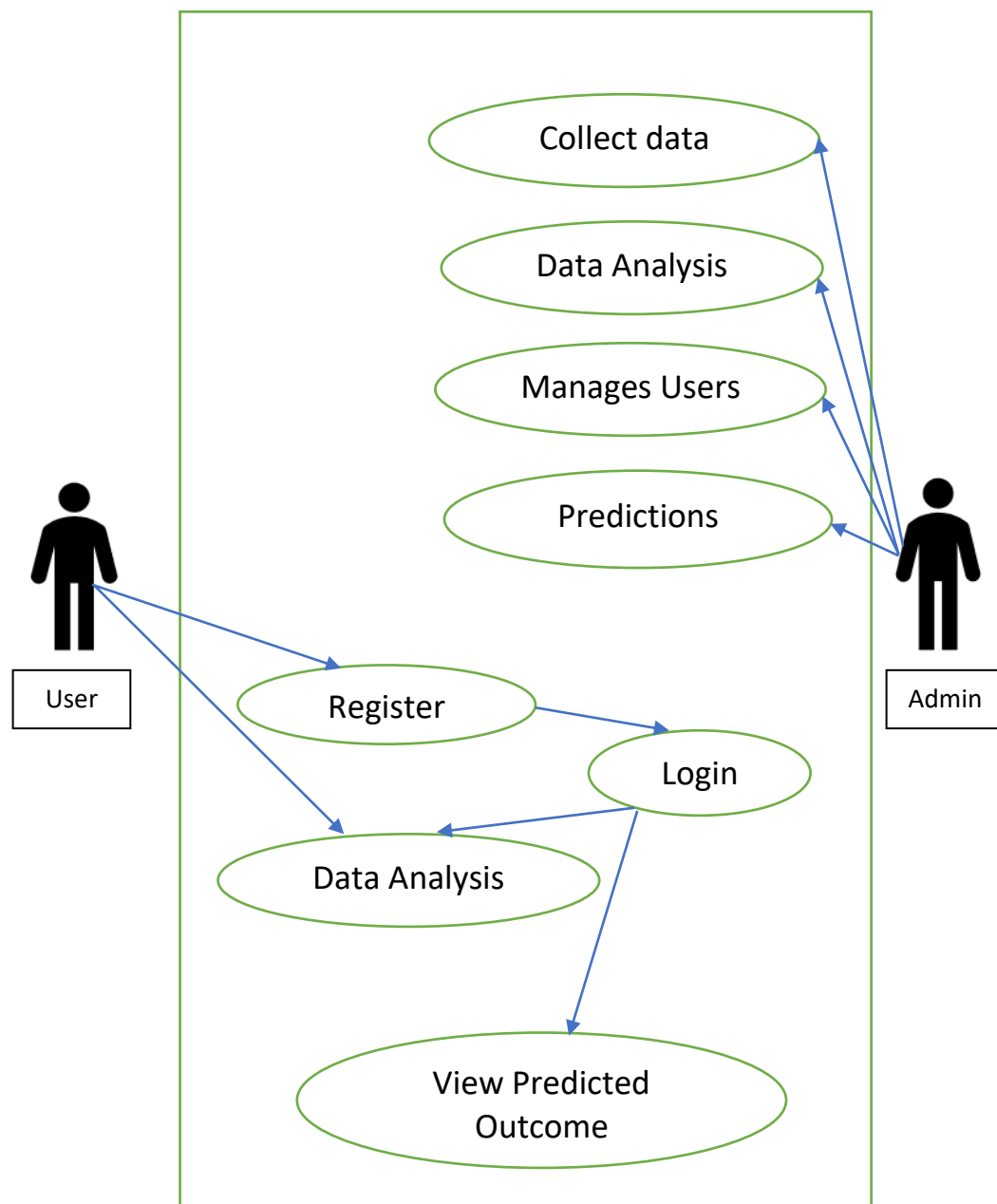


Figure 5.3.1: Use Case Diagram of System

5.3.2 System Flow Chart

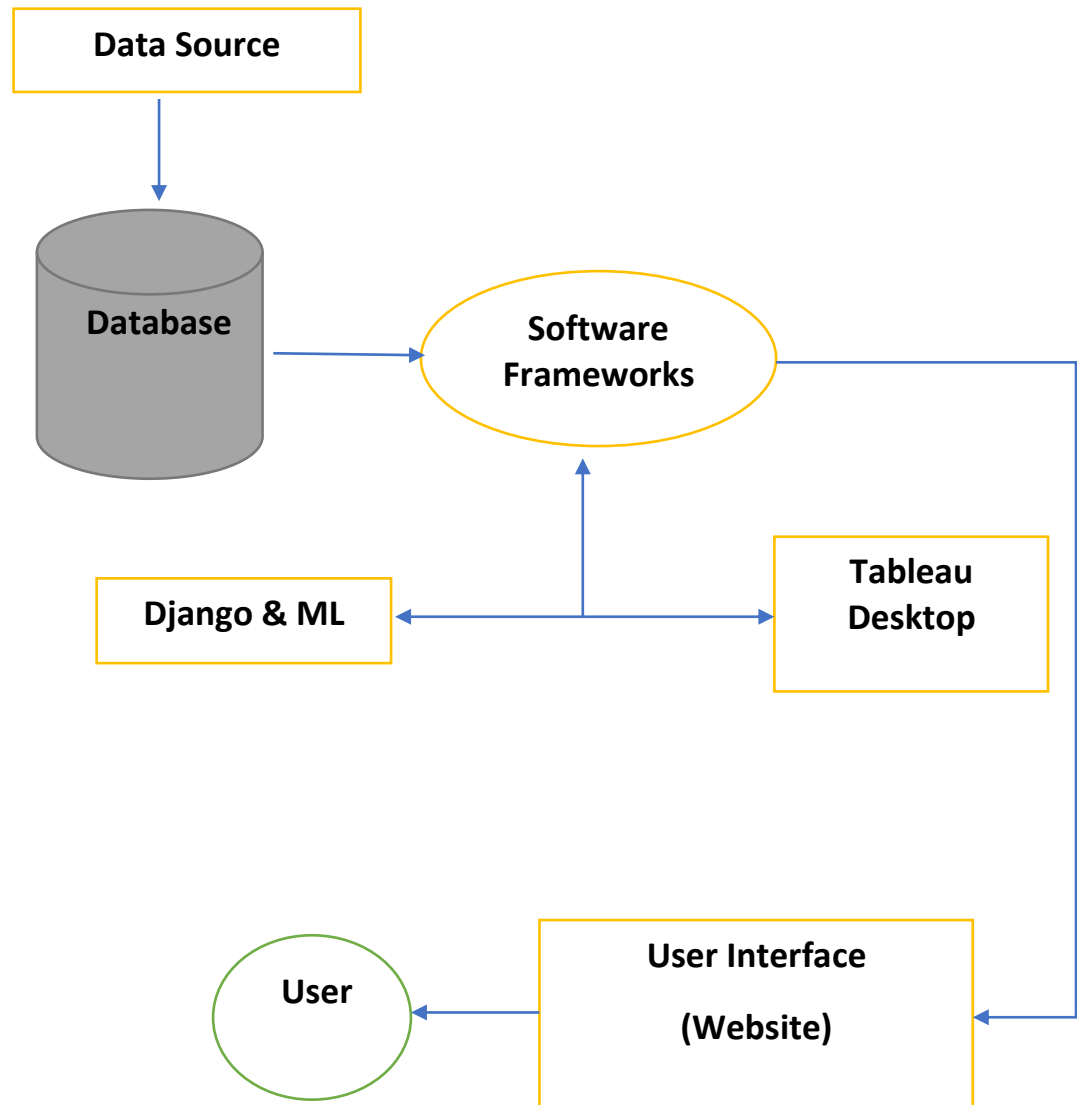


Figure 5.3.2: System Flow Chart Diagram

5.4 Pseudo code

```
from django.contrib.auth import logout, authenticate, login
from django.contrib.auth.decorators import login_required
from django.contrib.auth.models import Group
from django.core.checks import messages
from django.shortcuts import render, redirect
import pickle
import joblib
import numpy as np
from email.message import EmailMessage

import smtplib
from email.mime.text import MIMEText
from email.mime.image import MIMEImage
from email.mime.multipart import MIMEMultipart
# Create your views here.
from calc.decorator import unauthenticated_user
from calc.forms import CreateUserForm
from calc.models import Customer

@login_required(login_url='signin')
def home(request):

    context={'name':'dear user'}
    return render(request,'accounts/winWT.html',context)

@login_required(login_url='signin')
def add(request):

    model=pickle.load(open('pred_win.pkl','rb'))
    number1=int(request.POST.get('num1'))
    number2=int(request.POST.get('num2'))
    number3=int(request.POST.get('num3'))
    number4=int(request.POST.get('num4'))
    number5=int(request.POST.get('num5'))

    test=[number1,number2,number3,number4,number5]
```



```

test=np.asarray(test)
test=test.reshape(1,-1)
result=model.predict(test)
result=result[0]
if result:
    result="HOME TEAM WIL WIN THE MATCH"
else:
    result="HOME TEAM WILL LOSE THE MATCH"

    return render(request, 'accounts/winWT.html', {'result': result})

@login_required(login_url='signin')
def home2(request):

    context={'name':'dear user'}
    return render(request,'accounts/winWOT.html',context)

@login_required(login_url='signin')
def add2(request):

    model=pickle.load(open('pred_win_noToss.pkl','rb'))
    number1=int(request.POST.get('num1'))
    number2=int(request.POST.get('num2'))
    number3=int(request.POST.get('num3'))
    number4=int(request.POST.get('num4'))

    test=[number1,number2,number3,number4]
    test=np.asarray(test)
    test=test.reshape(1,-1)
    result=model.predict(test)
    result=result[0]

    if result:
        result="HOME TEAM WIL WIN THE MATCH"
    else:
        result="HOME TEAM WILL LOSE THE MATCH"

```

```

        return render(request, 'accounts/winWOT.html', {'result': result})

@login_required(login_url='signin')
def home3(request):

    context={'name':'dear user'}
    return render(request,'accounts/score.html',context)

@login_required(login_url='signin')
def add3(request):
    print("test")

    model=joblib.load(open('tree_model.pkl','rb'))
    print("test")

    batting_team=str(request.POST.get('num1'))
    bowling_team=str(request.POST.get('num2'))
    number3=int(request.POST.get('num3'))
    number4=int(request.POST.get('num4'))
    number5=int(request.POST.get('num5'))
    number6=int(request.POST.get('num6'))
    number7=int(request.POST.get('num7'))
    prediction_array=[]
    if batting_team == 'Chennai Super Kings':
        prediction_array = prediction_array + [1,0,0,0,0,0,0,0]
    elif batting_team == 'Delhi Daredevils':
        prediction_array = prediction_array + [0,1,0,0,0,0,0,0]
    elif batting_team == 'Kings XI Punjab':
        prediction_array = prediction_array + [0,0,1,0,0,0,0,0]
    elif batting_team == 'Kolkata Knight Riders':
        prediction_array = prediction_array + [0,0,0,1,0,0,0,0]
    elif batting_team == 'Mumbai Indians':
        prediction_array = prediction_array + [0,0,0,0,1,0,0,0]
    elif batting_team == 'Rajasthan Royals':
        prediction_array = prediction_array + [0,0,0,0,0,1,0,0]
    elif batting_team == 'Royal Challengers Bangalore':
        prediction_array = prediction_array + [0,0,0,0,0,0,1,0]
    elif batting_team == 'Sunrisers Hyderabad':
        prediction_array = prediction_array + [0,0,0,0,0,0,0,1]
    # Bowling Team
    if bowling_team == 'Chennai Super Kings':
        prediction_array = prediction_array + [1,0,0,0,0,0,0,0]
    elif bowling_team == 'Delhi Daredevils':

```

```

        prediction_array = prediction_array + [0,1,0,0,0,0,0,0]
    elif bowling_team == 'Kings XI Punjab':
        prediction_array = prediction_array + [0,0,1,0,0,0,0,0]
    elif bowling_team == 'Kolkata Knight Riders':
        prediction_array = prediction_array + [0,0,0,1,0,0,0,0]
    elif bowling_team == 'Mumbai Indians':
        prediction_array = prediction_array + [0,0,0,0,1,0,0,0]
    elif bowling_team == 'Rajasthan Royals':
        prediction_array = prediction_array + [0,0,0,0,0,1,0,0]
    elif bowling_team == 'Royal Challengers Bangalore':
        prediction_array = prediction_array + [0,0,0,0,0,0,1,0]
    elif bowling_team == 'Sunrisers Hyderabad':
        prediction_array = prediction_array + [0,0,0,0,0,0,0,1]

    prediction_array = prediction_array + [number3, number4, number5, number6, number7]
    print(prediction_array)

    print("-----")
    prediction_array = np.array([prediction_array])
    pred = model.predict(prediction_array)
    result = ""

    result= "The Predicted score is : {}".format(pred[0])
    print(result)
    print(type(pred))
    return render(request, 'accounts/score.html', {'result': result})

def signout(request):
    logout(request)
    return redirect('signup')

@login_required(login_url='signin')
def home4(request):

```

```

        context={'name':'dear user'}
        return render(request,'accounts/mom.html',context)

@login_required(login_url='signin')
def add4(request):

    team1=request.POST.get('num1')
    team2=request.POST.get('num2')
    if team2=='csk':
        result = 'Suresh Raina'
    elif team2=='srh':
        result = 'David Warner'
    elif team2=='mi':
        result = 'Rohit Sharma'
    elif team2=='rcb':
        result = 'Virat Kohli'
    elif team2=='dc':
        result = 'Rishab Pant'
    elif team2=='pbks':
        result = 'KL Rahul'
    elif team2=='RR':
        result = 'Sanju Samson'
    elif team2=='kkr':
        result = 'Andre Russell'

    return render(request, 'accounts/mom.html', {'result': result})

@login_required(login_url='signin')
def home5(request):

    context={'name':'dear user'}
    return render(request,'accounts/feedback.html',context)

@login_required(login_url='signin')
def add5(request):

    email=request.POST.get('num1')
    feed=request.POST.get('num2')

```

```

num=request.POST.get('num3')

msg = MIMEMultipart()
msg['Subject'] = 'You Got A Feed back from '+email
msg['From'] = 'editorial.iplprediction@gmail.com'
msg['To'] = 'editorial.iplprediction@gmail.com'

text = MIMEText("Feed: "+feed + "\n \n contact number:" + num + "\n Email: " + email)
msg.attach(text)
#image = MIMEImage(img_data, name=os.path.basename(ImgFileName))
#msg.attach(image)

s = smtplib.SMTP('smtp.gmail.com', 587)
s.ehlo()
s.starttls()
s.ehlo()
s.login('editorial.iplprediction@gmail.com', 'IPL@Predict')
s.sendmail(msg['From'], msg['To'], msg.as_string())
s.quit()

result = 'Feedback submitted successfully'

return render(request, 'accounts/feedback.html', {'result': result})

@unauthenticated_user
def singup(request):
    formSet=CreateUserForm()
    if request.method=="POST":
        formSet=CreateUserForm(request.POST)
        if formSet.is_valid():
            user=formSet.save()
            group=Group.objects.get(name='customer')
            user.groups.add(group)
            Customer.objects.create(
                user=user,
                name=user.username,
                email=user.email
            )
            return redirect('signin')
    context={'formSet':formSet}
    return render(request,'accounts/signup.html',context)

```

```

@unauthenticated_user
def signin(request):

    if request.method == "POST":
        username=request.POST.get('username')
        password=request.POST.get('password')
        user=authenticate(request,username=username,password=password)
        if user is not None:
            login(request,user)
            return redirect("winWT")
        else:
            context = {'message':'wrong pass or username'}
            return render(request, 'accounts/signin.html', context)

context={}
return render(request,'accounts/signin.html',context)

```

The code given above is used to deploy the ML models as pickles and to request data from user when registering and check the data when logging whether it is correct or wrong. And also, we have given feedback for users where this parameter works with smtpplib and MIME. It takes and sends the feedback to the email which was given in backend with the help of inbuilt HTTP in Django.

6. Testing

6.1 Functional Testing

Functional testing is the process through which QA 's determines the, if a piece of a software is acting accordance with the pre-determined requirements. There are different types of functional testing.

6.2 Structural Testing

Structural testing also known as the glass box testing or white box testing. It is the type of testing is used for the testing the structure of the code.

There are two types of structural testing. They are:

6.2.1 White Box Testing

It is a software testing technique in which tested to verify the flow of input-output and to improve design, usability and security.in white box testing code is visibility to user. So, it is also called clear box testing.

6.2.2 Black Box Testing:

It is a software testing technique in which the functionalities of the software applications are tested without the knowledge of internal code structure, implementation details and internal paths. It is also called behavioral testing.

6.3 Levels of Testing

There are four levels of testing. They are

6.3.1 Unit Testing

Unit testing is the basic or first level of testing part. In this part individual units or components are tested. This testing is done by developer. It is the developer's duty to make sure that the application is as per the design which was planned. Each phase of the project design and coding has been thoroughly tested. The testing of the interface done to make sure that the flow of information into and out of the program is completely as per planned and proper. The temporarily generated output data is made sure to maintain its integrity throughout the algorithm's execution. Finally, all error handling ways are also tested.

6.3.2 Integration Testing

Integration testing is a type of testing that comes next of unit testing in these testing. Software modules are integrated logically and tested as a group. The typical software consists of multiple software modules coded by different programmers. The main purpose of this testing is to find the faults in integrated modules. With these integration testing, we can resolve issues generated while module integration. During this phase, we don't emphasis on functionality of every module. We tested the overall model in this level.

6.3.3 System Testing

It is the testing in which the complete application is tested as whole. system testing is undertaken by the group independent testers who are not the part of developing the program.

6.3.4 Acceptance Testing

It is final level of testing. It is conduct to determine whether the system is ready to release or not.

6.4 Verification and Validation

The testing procedure is a part of the overall topic of verification and validation. We must accept the device specifications and make every effort to satisfy the customer's requirements, and we must check and test the product for this sole purpose. Validation and verification are completely two different parts. One is done to ensure that the program is correctly incorporates a particular functionality, and the other is done to ensure that the final product meets the customer requirements.

Verification of the project is to confirm that the project has met all the required specifications. We made sure that our project is up to the standard as we have planned at the beginning of our project development. It is said to be static testing.

Validation of the project is to check whether the model is up to the standards. It is the dynamic testing. We also include the code execution in this phase.

7. Project Results

7.1 Data Analysis

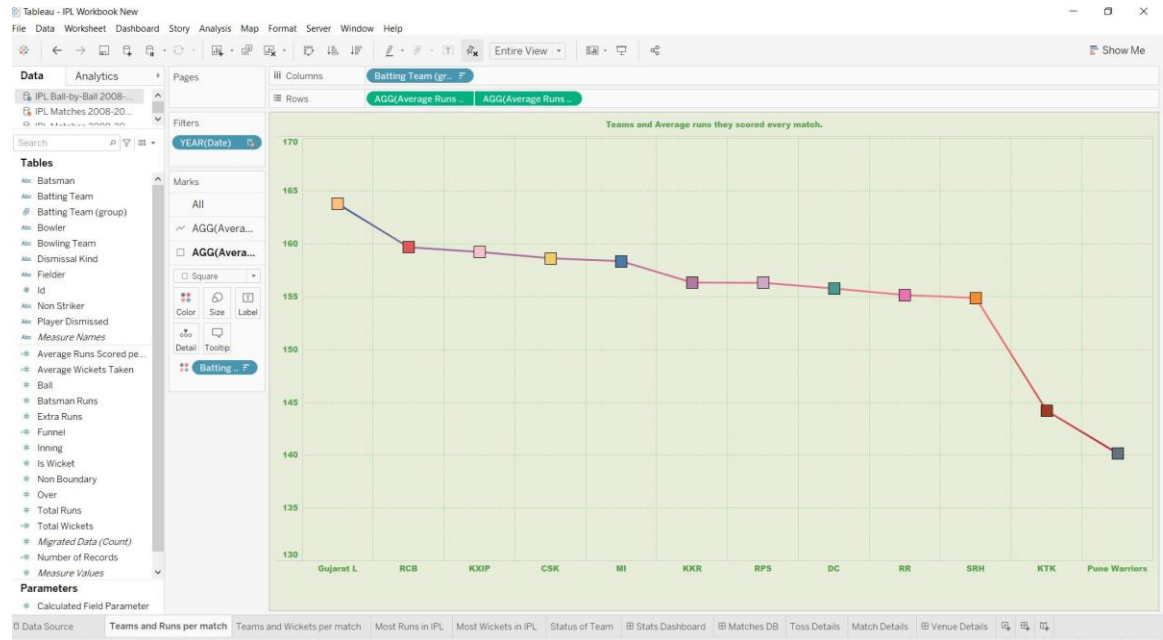


Figure 7.1(a) Average runs scored by each team in IPL

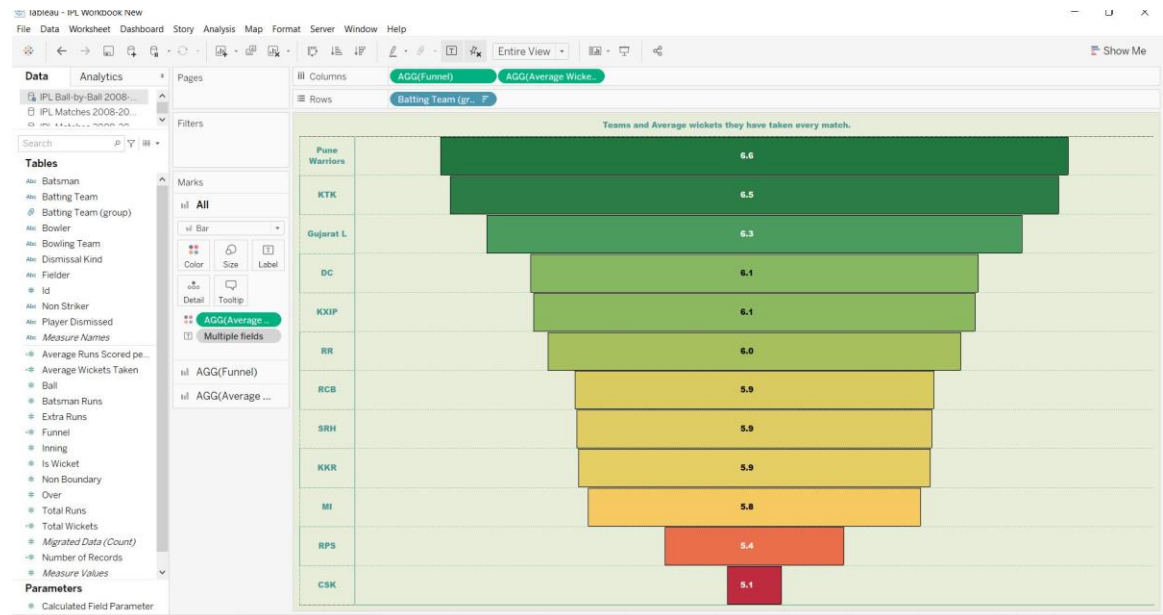


Figure 7.1(b) Average wickets taken by each team in IPL.

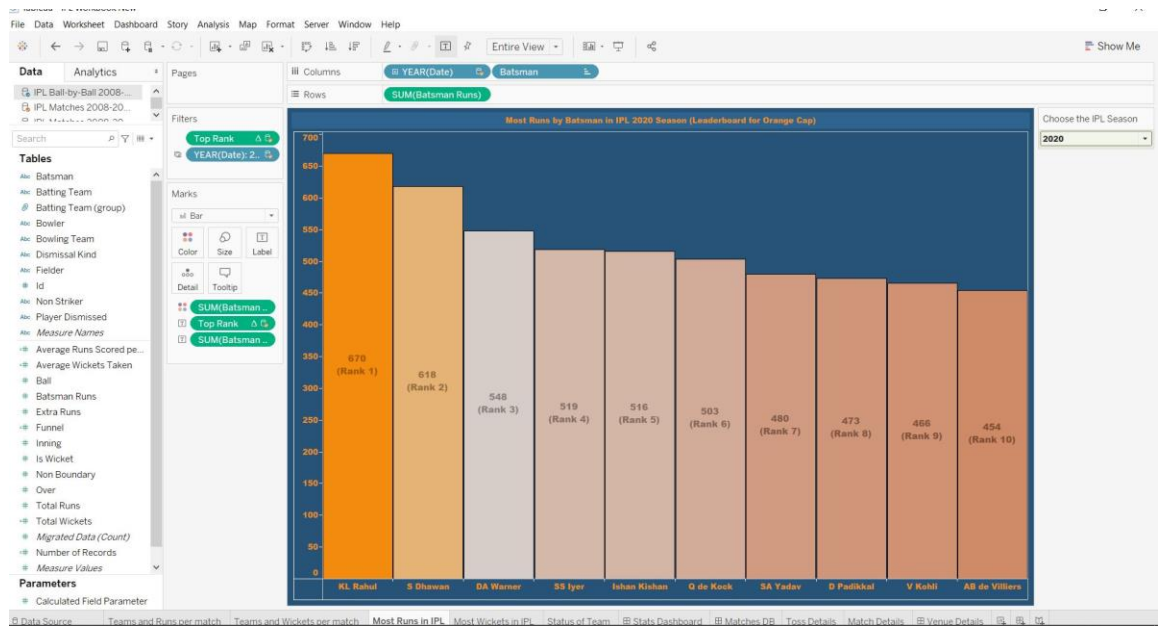


Figure 7.1(c) Orange cap holder of every IPL season.

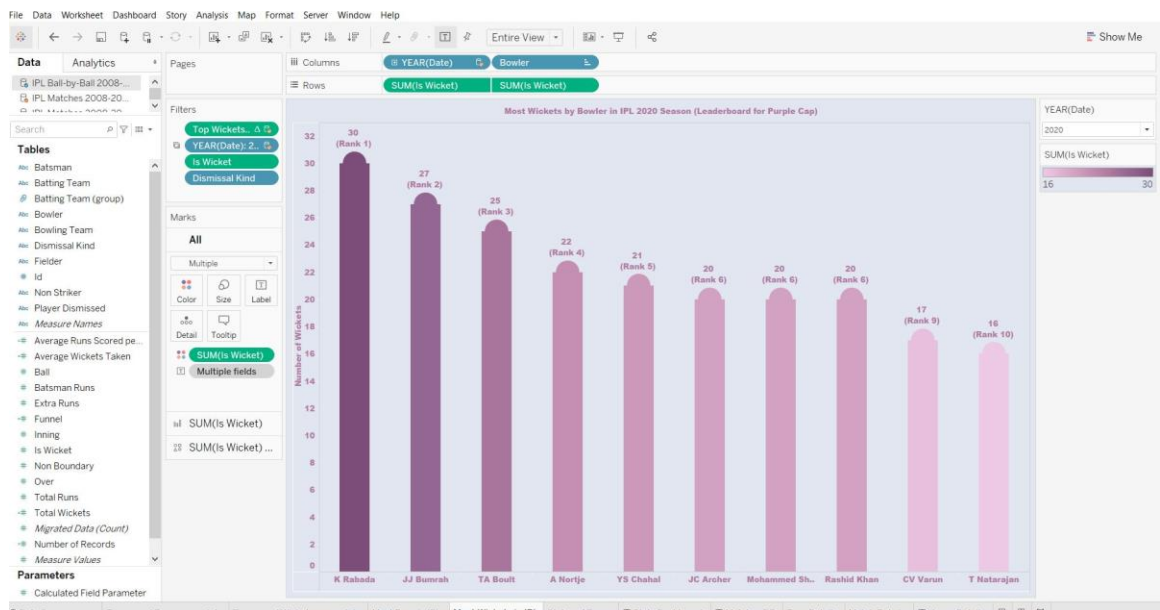


Figure 7.1(d) Purple cap holder of every IPL season.

7.2 Dash Boards

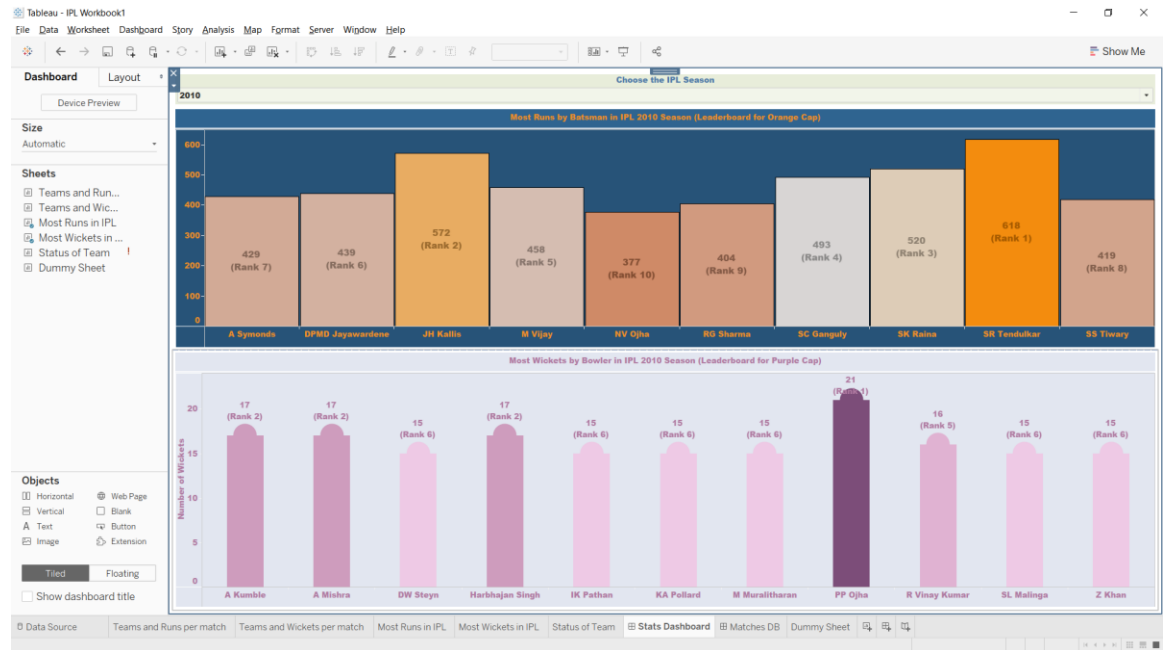


Figure 7.2(a) Stats Dash Board

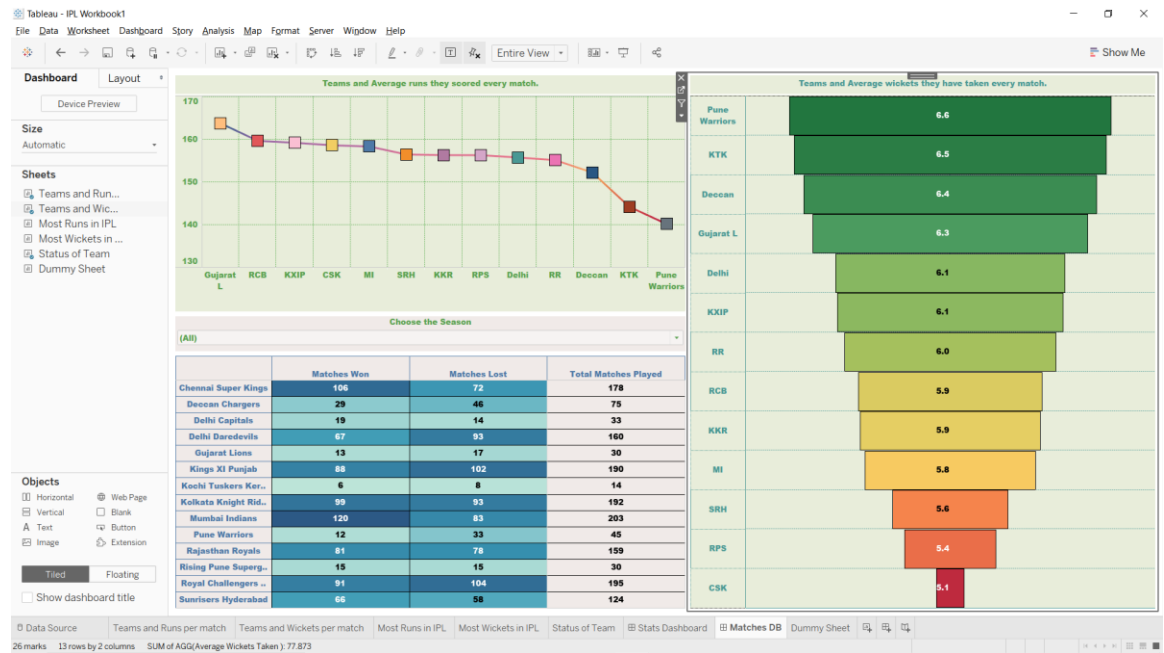


Figure 7.2(b) Matches Dash Boards

7.3 Algorithms

Ensemble Technique Algorithm

It is an Algorithm which combines all remaining algorithms and gives a precised accurate output. There are the algorithms which are used in our project,

7.3.1 Linear regression

The fitting of a linear function of one or more inputs to an output is known as linear regression. The fitting of a straight line with input x and output y on the form $y = w_1x + w_0$, where w_0 and w_1 are real-valued coefficients to be known, is the univariate case. The most popular method for determining the weights in a linear regression problem is to minimise the squared loss function. The next step is to determine the weight vector w . We will find a special global minimum if we choose the weights in this manner.

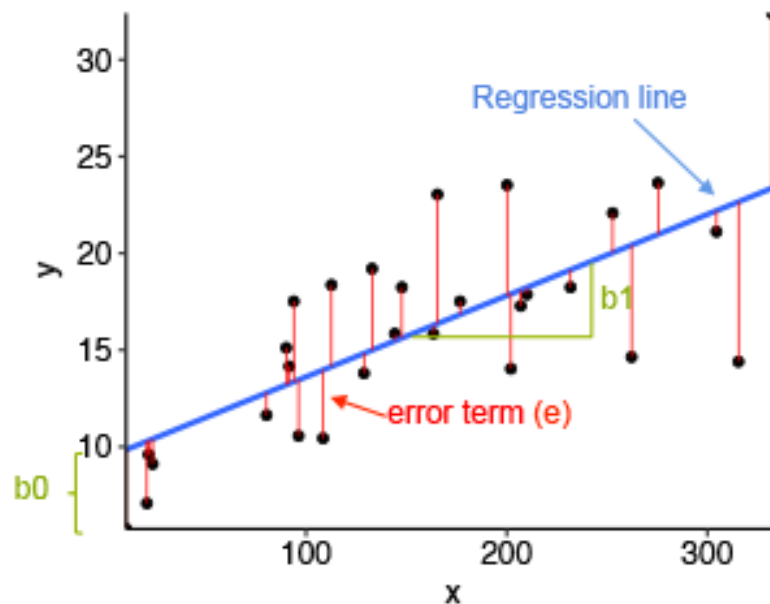


Figure 7.3.1 Linear regression

To prepare or train the linear regression equation from data, various techniques can be used, the most popular of which is called Ordinary Least Squares. Ordinary Least

Squares Linear Regression, or simply Least Squares Regression, is a term used to describe a model created in this manner.

7.3.2 Neural Networks

A node layer contains an input layer, one or more hidden layers, and an output layer in artificial neural networks (ANNs). Each node, or artificial neuron, is connected to the others and has a weight and threshold associated with it. If a node's performance exceeds a certain threshold value, the node is enabled, and data is sent to the next layer of the network. Otherwise, no data is passed on to the network's next layer.

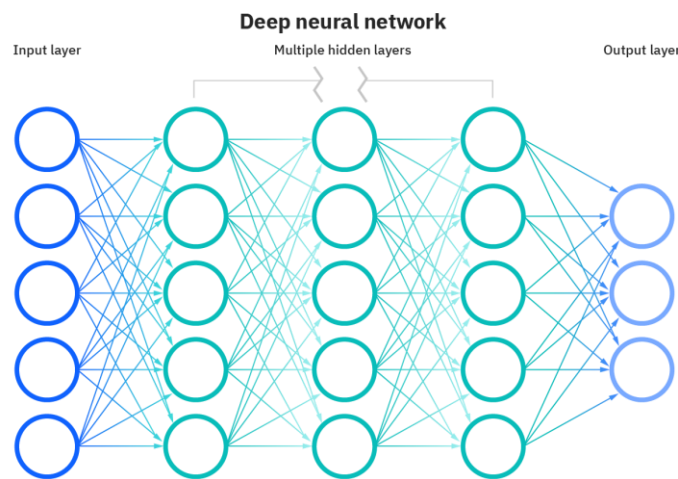


Figure 7.3.2 Deep Neural Network

Training data is used by neural networks to learn and enhance their performance over time. However, once these learning algorithms have been fine-tuned for precision, they become powerful tools in computer science and artificial intelligence, enabling us to quickly classify and cluster data. When compared to manual identification by human experts, tasks in speech recognition or image recognition will take minutes rather than hours. Google's search algorithm is one of the most well-known neural networks.

7.3.3 Decision tree

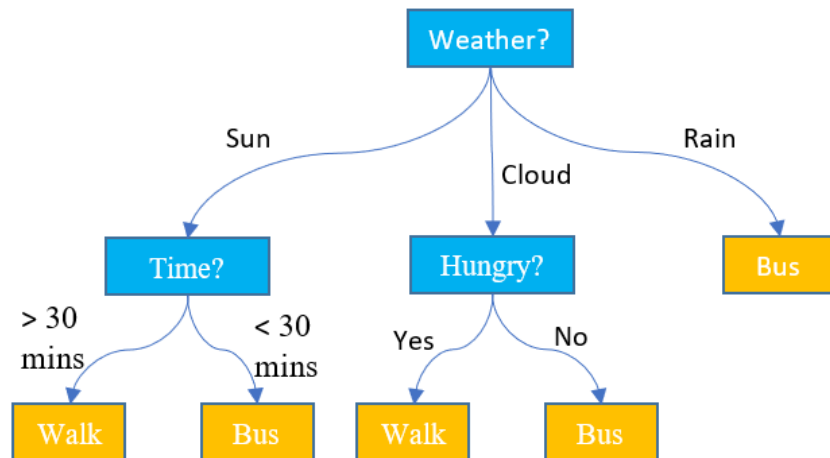


Figure 7.3.3(a) Working of Decision Tree

Each category is divided based on the probability of the effect on dependent variable (fuel consumption). Variable having highest probability is placed at top and others are placed according to their respective chances.

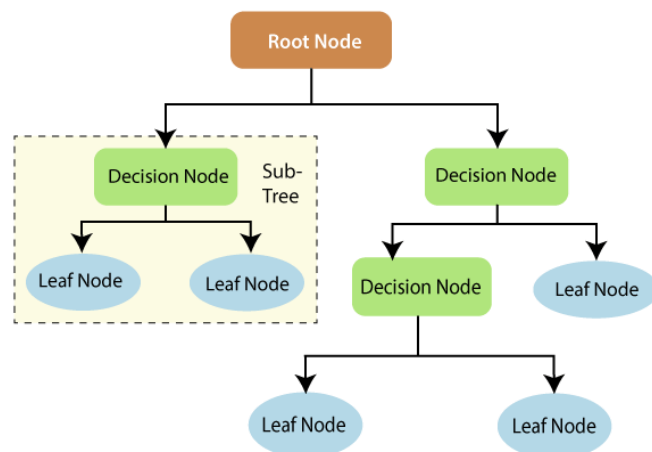


Figure 7.3.3(b) Leveling of Nodes

Regression Trees

Regression trees are decision trees used for regression, that is their target variable can take continuous values. A regression tree is a tree of nodes where each leaf node has a linear function of some subset of numerical attributes, rather than a single value which is the case for classification trees. For example, a regression tree for fuel prediction may have leaf nodes that contain linear functions of vehicle weight, road slope and engine strength. The learning algorithm must decide when to stop splitting and start to apply linear regression over the attributes. An example of a regression tree is illustrated in Figure 7.3.3(c). In order to place the nodes and which node to choose as the root is decided by examining the entropy and information gain of the attributes. Information gain is the expected reduction of entropy achieved after eliminating an attribute from the equation.

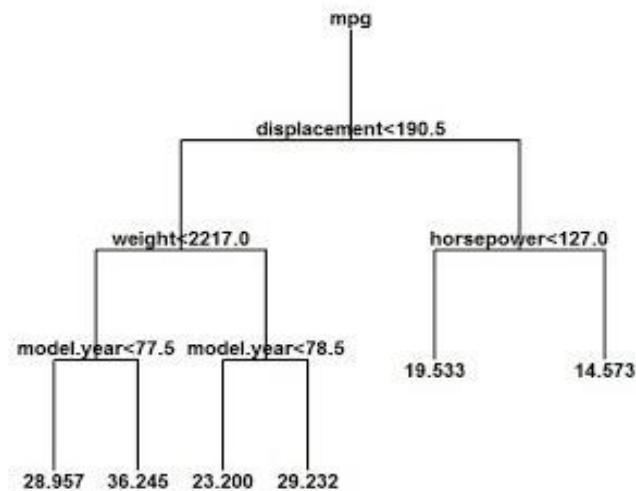


Figure 7.3.3(c) Example of Regression Tree

7.3.4 Random Forest Algorithm

It is an extension of bootstrap aggregation (bagging) of decision trees and can be used for classification and regression problems.

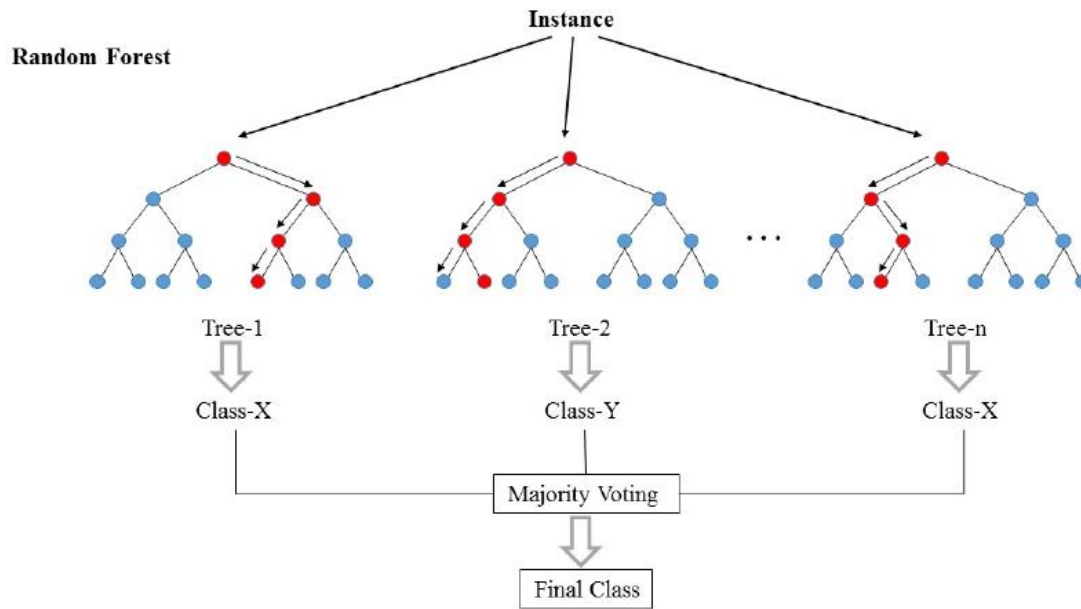


Figure 7.3.4(a) Random Forest

We can see how an example is categorized using n trees in the figure above, with the final prediction made by taking a vote from all n trees.

RFs are also used as an ensemble or bagging tool of machine learning. Random forest is a supervised learning algorithm in essence. This can be applied to both regression and classification problems. However, since it is more intuitive and easier to understand, we will explore its use for classification. Because of its simplicity and stability, random forest is one of the most widely used algorithms. Like bagging, random forest entails building a large number of decision trees from bootstrap samples from the training dataset.

Bagging and boosting

- Bagging is a technique for combining predictions of the same kind. Boosting is a technique for combining various forms of predictions.
- Bagging addresses over-fitting problems in a model by reducing uncertainty rather than bias. Boosting reduces bias rather than variation.
- In Bagging, each model is given the same amount of weight. Models are weighed in Boosting based on their efficiency.

- In Bagging, models are produced separately. The success of a previously developed model in Boosting has an effect on new models.

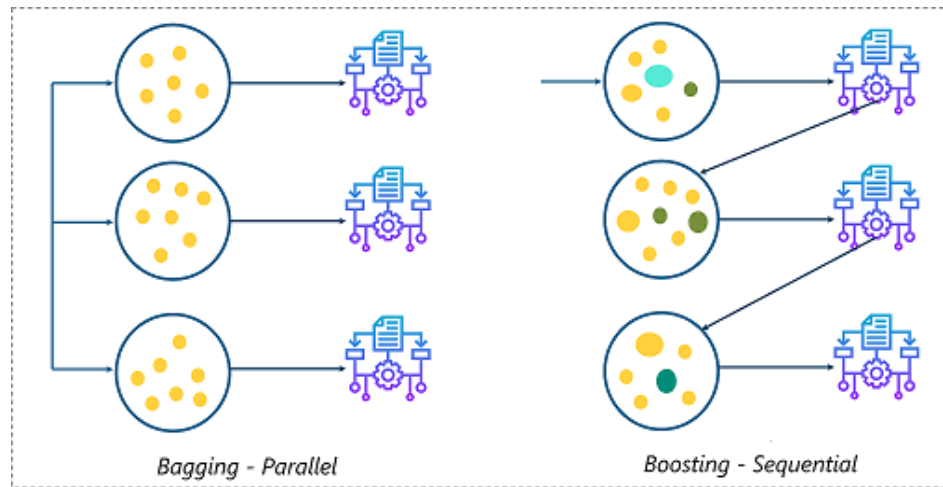


Figure 7.3.4(b) Bagging and Boosting

Bootstrapping is a sampling method that involves replacing subsets of observations from the actual given dataset. The size of the subsets is equal to the actual set.

Similarities Between Bagging and Boosting –

1. Both are ensemble methods to get N learners from 1 learner.
2. Both generate several training data sets by random sampling.
3. Both make the final decision by averaging the N learners (or taking the majority of them i.e., Majority Voting).
4. Both are good at reducing variance and provide higher stability.

7.3.5 Logistic Function

The logistic equation, which is at the core of the system, is called logistic regression. The logistic function, also known as the sigmoid function, was created by statisticians to explain the properties of population growth in ecology, such as how it rises rapidly and eventually reaches the environment's carrying capacity. It's an S-shaped

curve that can map any real-valued number to a value between 0 and 1, but never exactly between those two points.

Formula: $1 / (1 + e^{-\text{value}})$

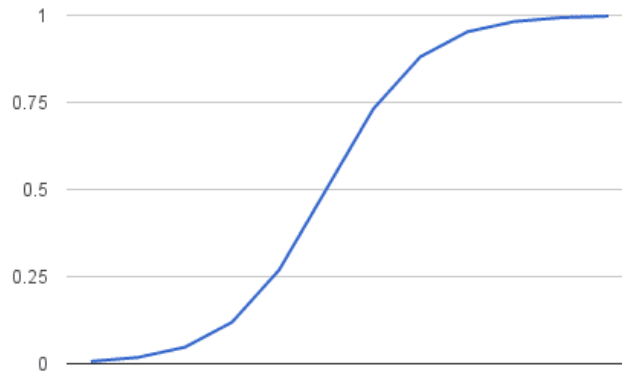


Figure 7.3.5 Logistic Function

Representation Used for Logistic Regression

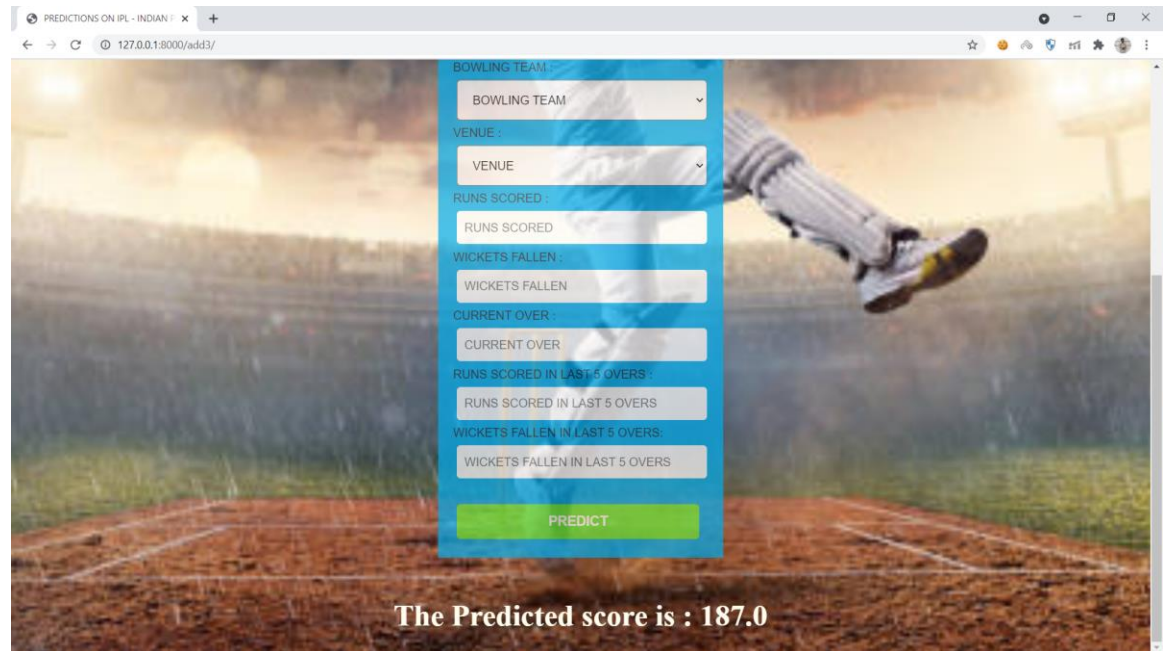
- Like linear regression, logistic regression uses an equation as a representation.
- To estimate an output value (y), input values (x) are combined linearly using weights or coefficient values (referred to as the Greek capital letter Beta). The output value being modelled is a binary value (0 or 1) rather than a numeric value, which is a crucial distinction from linear regression.

- The logistic regression equation:

$$y = e^{(b_0 + b_1 * x)} / (1 + e^{(b_0 + b_1 * x)})$$

- Where y denotes the expected output, b0 denotes the bias or intercept expression, and b1 denotes the coefficient for a single input value (x). The b coefficient (a constant real value) for each column in the input data must be learned from the training data.

7.4 Predictions



PREDICTIONS ON IPL - INDIAN

BOWLING TEAM :
BOWLING TEAM

VENUE :
VENUE

RUNS SCORED :
RUNS SCORED

WICKETS FALLEN :
WICKETS FALLEN

CURRENT OVER :
CURRENT OVER

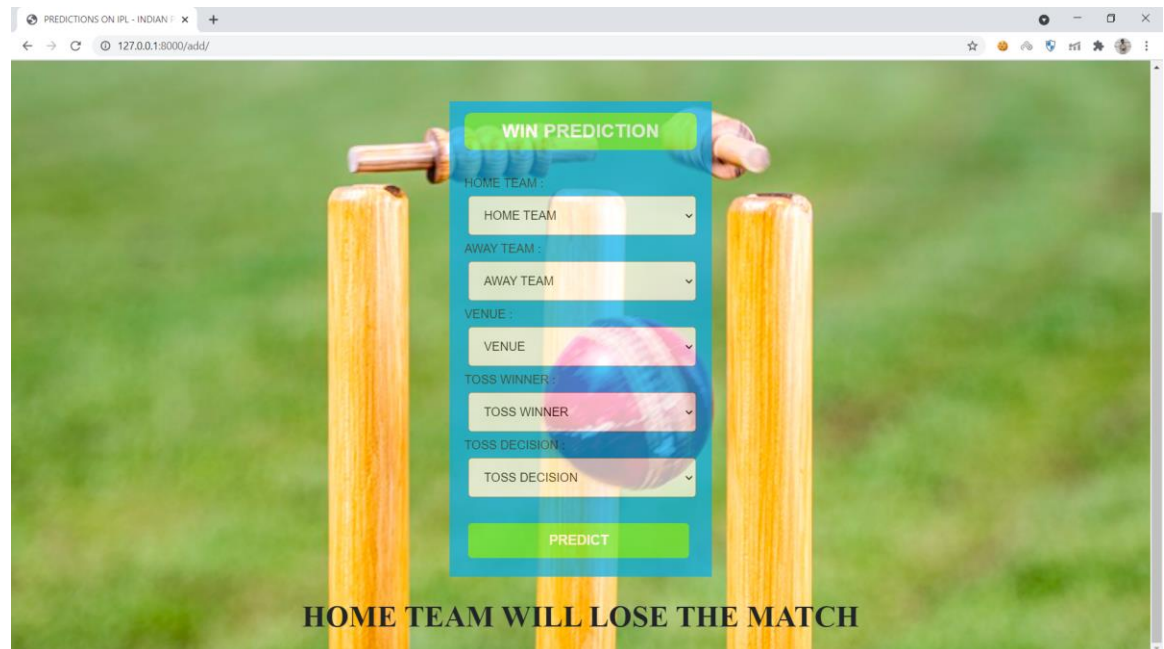
RUNS SCORED IN LAST 5 OVERS :
RUNS SCORED IN LAST 5 OVERS

WICKETS FALLEN IN LAST 5 OVERS :
WICKETS FALLEN IN LAST 5 OVERS

PREDICT

The Predicted score is : 187.0

Figure 7.3(a) Match Score Prediction.



PREDICTIONS ON IPL - INDIAN

WIN PREDICTION

HOME TEAM :
HOME TEAM

AWAY TEAM :
AWAY TEAM

VENUE :
VENUE

TOSS WINNER :
TOSS WINNER

TOSS DECISION :
TOSS DECISION

PREDICT

HOME TEAM WILL LOSE THE MATCH

Figure 7.3(b) Match Win Prediction.

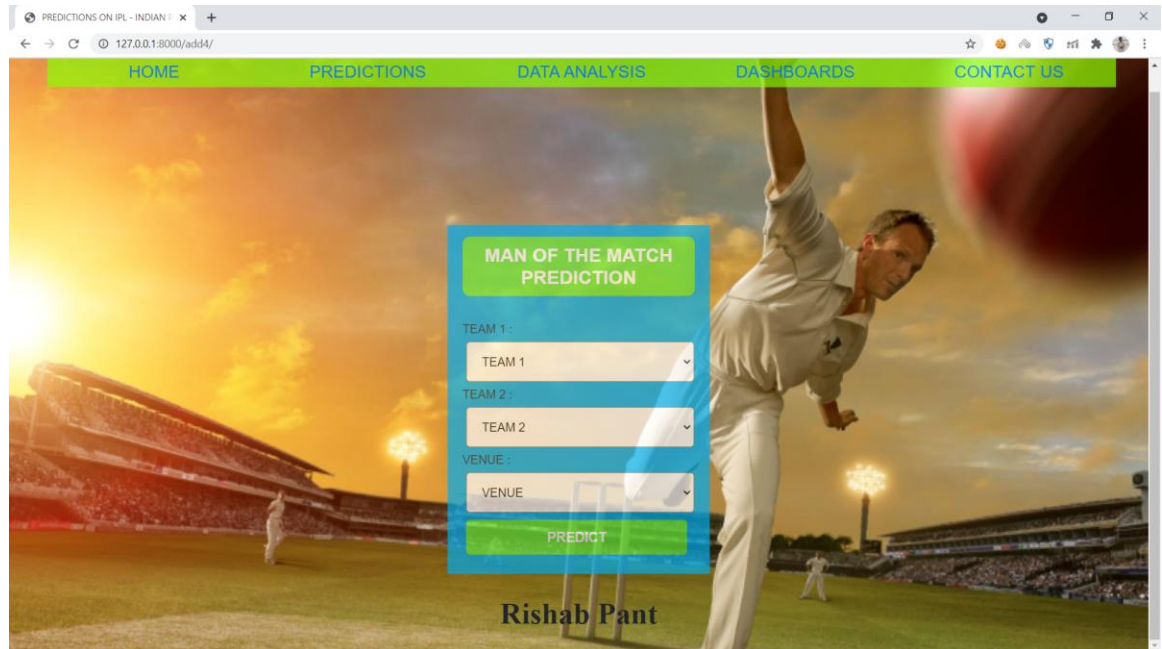


Figure 7.3(c) Prediction of Man of the Match.

7.5 Web Interface

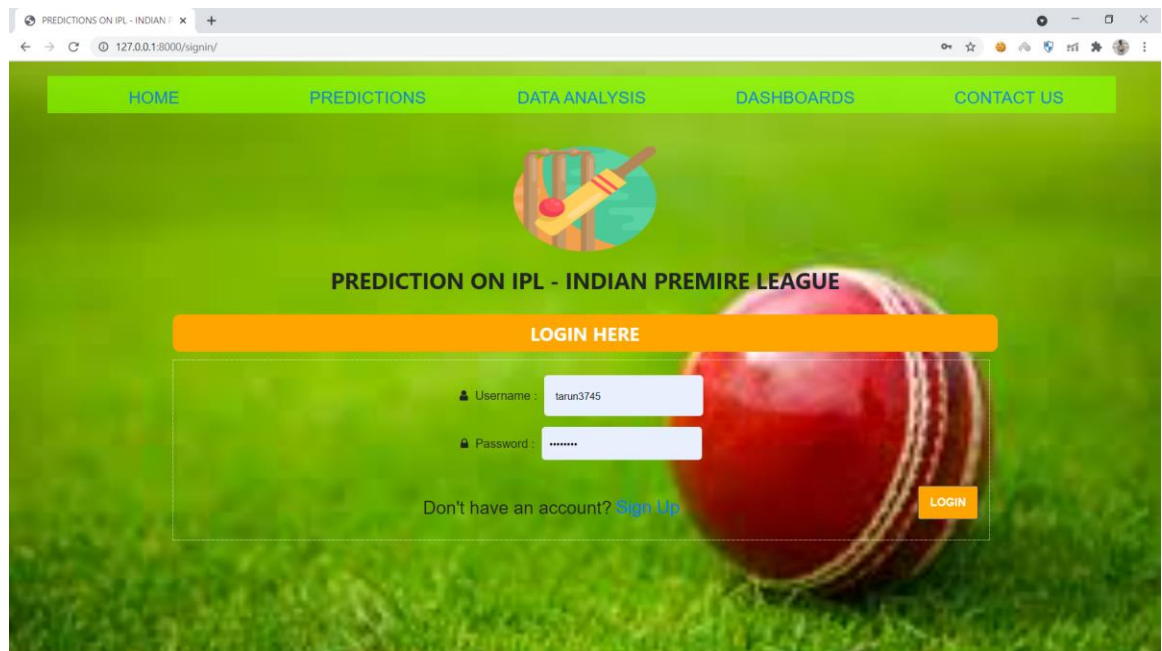


Figure 7.4(a) Login Page.

PREDICTIONS ON IPL - INDIAN

HOME PREDICTIONS DATA ANALYSIS DASHBOARDS CONTACT US

NEW USER

Username :

Email :

Password :

Confirm Password :

Re-enter Password :

Submit Now

Cancel

Already have an account? [Sign In](#)

Figure 7.4(b) New User Registration Page.

PREDICTIONS ON IPL - INDIAN

HOME PREDICTIONS DATA ANALYSIS DASHBOARDS CONTACT US

FEEDBACK FORM

Please enter your Email:

Please enter your Phone Number:

Feedback:
Enter your feedback here

SUBMIT

Figure 7.4(c) Feedback Page.

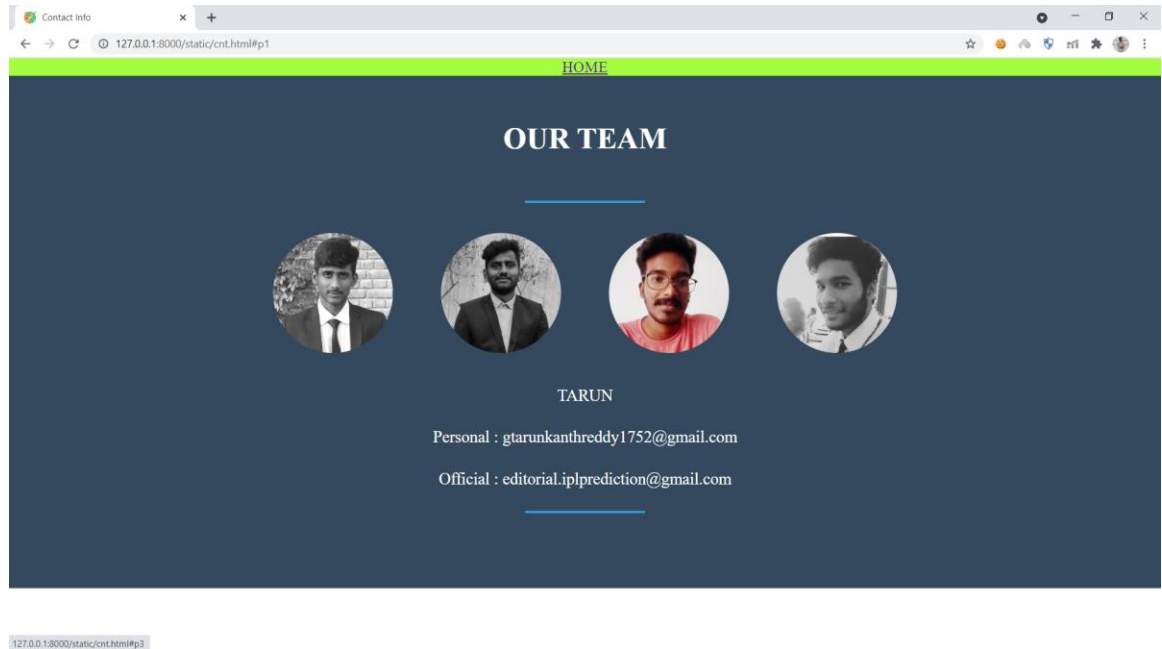


Figure 7.4(d) Contact Us Page.

8. Implementation

IPL winning prediction on the basis of previous years data. There are 2 categories in winning prediction 1) Before toss ,2) after toss. And it calculates the most probable winning team on the basis of their previous records of their players in that ground. Previous records Includes batsmen average and strike rate and bowler's economy. It takes the input of current run rate, current running over and current score to predict the projected score and it gives the projected score as output. Finally on the basis of these all inputs it gives the most probable winning team as output.

The data is analyzed to find correlations between different features. A correlation plot is presented below. Strong correlations are found between the features describing the engine characteristics. One interesting question is if the engine features could be reduced to one or two descriptive features using principal component analysis (pca) or some other method for dimensionality reduction. This question is beyond the scope of this study but could be worth investigating in future research.

```

1 # Information (not-null count and data type) About Each Column
2 data.info()

<class 'pandas.core.frame.DataFrame'>
RangeIndex: 76014 entries, 0 to 76013
Data columns (total 15 columns):
#   Column                Non-Null Count  Dtype
---  -
0   mid                    76014 non-null  int64
1   date                   76014 non-null  object
2   venue                  76014 non-null  object
3   batting_team           76014 non-null  object
4   bowling_team           76014 non-null  object
5   batsman                76014 non-null  object
6   bowler                 76014 non-null  object
7   runs                   76014 non-null  int64
8   wickets                76014 non-null  int64
9   overs                  76014 non-null  float64
10  runs_last_5            76014 non-null  int64
11  wickets_last_5         76014 non-null  int64
12  striker                76014 non-null  int64
13  non-striker            76014 non-null  int64
14  total                  76014 non-null  int64
dtypes: float64(1), int64(8), object(6)
memory usage: 8.7+ MB

```

Figure 8(a) Info of the Data

```

1 # Describing Numerical Values of the Dataset
2 data.describe()

```

	mid	runs	wickets	overs	runs_last_5	wickets_last_5	striker	non-striker	total
count	76014.000000	76014.000000	76014.000000	76014.000000	76014.000000	76014.000000	76014.000000	76014.000000	76014.000000
mean	308.627740	74.889349	2.415844	9.783068	33.216434	1.120307	24.962283	8.869287	160.901452
std	178.156878	48.823327	2.015207	5.772587	14.914174	1.053343	20.079752	10.795742	29.246231
min	1.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	0.000000	67.000000
25%	154.000000	34.000000	1.000000	4.600000	24.000000	0.000000	10.000000	1.000000	142.000000
50%	308.000000	70.000000	2.000000	9.600000	34.000000	1.000000	20.000000	5.000000	162.000000
75%	463.000000	111.000000	4.000000	14.600000	43.000000	2.000000	35.000000	13.000000	181.000000
max	617.000000	263.000000	10.000000	19.600000	113.000000	7.000000	175.000000	109.000000	263.000000

Figure 8(b) Describing Numerical Values

```

1 # Number of Unique Values in each column
2 data.nunique()

mid          617
date         442
venue        35
batting_team 14
bowling_team 14
batsman       411
bowler        329
runs          252
wickets       11
overs         140
runs_last_5   102
wickets_last_5 8
striker       155
non-striker   88
total        138
dtype: int64

```

Figure 8(c) Unique Values in each and every column and row

With all the algorithms developed in the project are then calculated for the accuracy and all the accuracies of algorithms are shown in a bar plot with the help of seaborn attribute as shown in Figure 8(d).

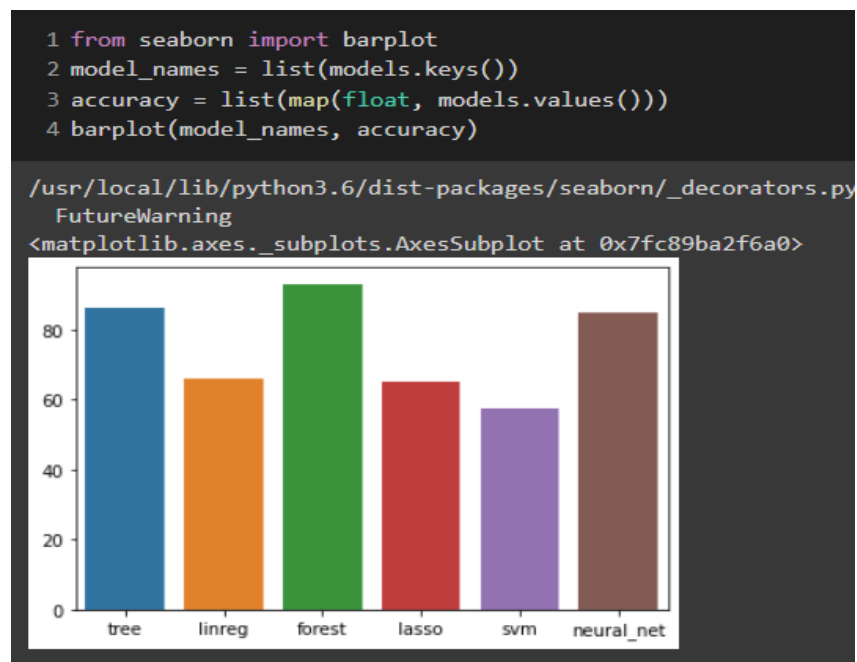


Figure 8 (d) Model Performance Comparison

9. Conclusion

This project is designed to implement as a system that will be useful for the prediction and previous analysis of ipl matches. IPL predictor offers different types of predictions like toss prediction, score prediction, win prediction to the user, where some of them are in our project and remaining in further can be developed on the basis of our project. So, user can choose which option he wants. We Implement Machine Learning models with ensemble techniques to get accurate prediction. Our analysis shows the significant correlation between different parameters.

The result obtained from our analysis is fairly accurate. The prediction result obtained is analyzed among various predictive algorithms and the best result is taken into consideration.

10. References

- 1) Linear Regression Algorithm,

Available from: <https://mylearningsinaiml.wordpress.com/ml-models/linear-regression/>

- 2) YouTube – A platform where all required fields are available,

<https://www.youtube.com/>

- 3) Wikipedia of IPL,

Available from: https://en.wikipedia.org/wiki/Indian_Premier_League

- 4) Neural Networks by IBM cloud,

Available from: <https://www.ibm.com/cloud/learn/neural-networks>

- 5) Random Forest Algorithm, Machine Learning,

Available form: <https://builtin.com/data-science/random-forest-algorithm>

- 6) How HTML and CSS works and a brief info on them,

Available from:

[https://www.w3.org/standards/webdesign/htmlcss#:~:text=HTML%20\(the%20Hyertext%20Markup%20Language,for%20a%20variety%20of%20devices.](https://www.w3.org/standards/webdesign/htmlcss#:~:text=HTML%20(the%20Hyertext%20Markup%20Language,for%20a%20variety%20of%20devices.)

7) Django backend framework in python by Greek for Greeks,

Available from: <https://www.geeksforgeeks.org/django-tutorial/>