A PROJECT REPORT ON
Cricket Score Predictor and Analysis

Submitted in partial fulfilment for Degree of

MASTER OF COMPUTER APPLICATION

By

Anil Rode

Suraj Sawant

Sushant Padale

Under the guidance of

Prof. Rahul Kotawadekar
(Department of MCA)

Submitted to
Department of MCA
FINOLEX ACADEMY OF MANAGEMENT AND TECHNOLOGY, RATNAGIRI

FINOLEX ACADEMY OF MANAGEMENT AND TECHNOLOGY RATNAGIR

This is to certify that the project report titled:

Cricket Score Predictor and Analysis

Carried out by student of MCA

Submitted By:

Anil Rode (45)

Suraj Sawant (48)

Sushant Padale (32)

In partial fulfilment of the award for degree of

MASTER OF COMPUTER APPLICATION

From Mumbai University

And are the bonafide records of the work done by them

during the Semester III of A.Y 2022-2023

Internal Guide   **HOD**

(Prof.Rahul Kotawadekar)   **(Prof. Tejas V.Joshi)**

CONTENTS OF SRS

ABSTRACT

In today's date data analysis is need for every data analytics to examine the sets of data to extract then useful information from it and to draw conclusion according to the information. Data analytics techniques and algorithms are more used by the commercial industries which enables them to take precise business decisions. It is also used by the analysts and the experts to authenticate or negate experimental layouts, assumptions and conclusions. In recent years the analytics is being used in the field of sports to predict and draw various insights. Due to the involvement of money, team spirit,city loyalty and a massive fan following, the outcome of matches is very important for all stake holders.

In this paper, the past seven year's data of IPL containing the player's details, match venue details, teams, ball to ball details, is taken and analyzed to draw various conclusions which help in the improvement of a player's performance. Various other features like how the venue or toss decision has influenced the winning of the match in last seven years are also predicted. Various machine learning and data extraction models are considered for prediction are Linear regression, Decision tree,K-means, Logistic Regression etc. The cross validation score and the accuracy are also calculated using various machine learning algorithms. Before prediction we have to explore and visualize the data because data exploration and visualization is an important stage of predictive modeling.

ACKNOWLEDGEMENT

The satisfaction and euphoria that accompany a successful completion of any task would be incomplete without the mention of people who made it possible, success is the epitome of hard work and perseverance, but steadfast of all is encouraging guidance. So, it is with gratitude that we acknowledge all those whose guidance and encouragement served as beacon of light and crowned our effort with success. We would like to thank Mr.Rahul Kotawadekar Sir, Finloex College, Ratanagiri, for providing an excellent academic environment in the college and his never-ending support for the MCA program.

We would like to express our gratitude towards Miss Harshada Salvi Meadum,and HOD, Department of Master Computer Application (MCA) Finloex College,Ratanagiri,who provided guidance and gave valuable suggestions regarding the project.We consider it a privilege and honour to express our sincere gratitude to our internal guide Mr. Rahul Kotawadekar, Master Computer Application for their valuable guidance throughout the tenure of this project work. We would also like to thank all the faculty members who have always been very Co-

operative and generous. Conclusively, we also thank all the non-teaching staff and all others who have done immense help directly or indirectly during our project.

Mr. Anil Rode

Mr. Suraj Sawant

Mr. Sushant Padale

Introduction

Machine Learning is a branch of Artificial Intelligence that aims at solving real-life engineering problems. This technique requires no programming, whereas it depends on only data learning where the machine learns from pre-existing data and predicts the result accordingly. Machine Learning methods have benefit of using decision trees, heuristic learning, knowledge acquisition, and mathematical models. It thus provides controllability, observability, stability and effectiveness. Cricket is being played in many countries around the world. There are a lot of domestic and international cricket tournaments being held in many countries.

The cricket game has various forms such as Test Matches, Twenty20 Internationals, Internationals one day, etc. IPL is also one of them, and has great popularity among them. It's a twenty-20 cricket game league played to inspire young and talented players in India. The league was conducted annually in March, April or May and has a huge fan base among India. There are eight teams which represent eight cities which are chosen from an auction. These teams compete against each other for the trophy. The whole match depends on the luck for the team, player's performance and lot more parameters that will be taken in to the consideration.

The match that is played before the day is also will make a change in the prediction. The stakeholders are much more benefited due to the huge popularity and the huge presence of people at the venue. The accuracy of a data depends on the size of the data we take for analysing and the records that are taken for predicting the outcome. Cricket is a game played between two teams comprising of 11 players in each team. The result is either a win, loss or a tie. However, sometimes due to bad weather conditions the game is also washed out as Cricket is a game which cannot be played in rain. Moreover, this game is also extremely unpredictable because at every stage of the game the momentum shifts to one of the teams between the two.

A lot of times the result gets decided on the last ball of the match where the game gets really close. Considering all these unpredictable scenarios of this unpredictable game, there is a huge interest among the spectators to do some prediction either at the start of the game or during the game. Many spectators also play betting games to win money.

Plan of Implementation

The project can be broken down into 7 main steps which are as follows:

1. Understand the dataset.
2. Clean the data.
3. Analyse the candidate columns to be Features.
4. Process the features as required by the model/algorithm.
5. Train the model/algorithm on training data.
6. Test the model/algorithm on testing data.
7. Tune the model/algorithm for higher accuracy.

Problem Statement

To predict the results of an IPL match using machine learning techniques or algorithms such as Logistic Regression, Gaussian Naive Bayes, K Nearest Neighbours, SVM, Gradient boost algorithm, Decision tree and Random forest.

We have used 17 features which are as follows: season, city, date, team1, team2, toss_winner, toss_decision, result, dl_applied, winner, win_by_runs, win_by_wickets, player_of_match, venue,umpire1, umpire2 and umpire3.

Objective of the Project

The main objective of this project is to give the team players information about how each venue makes a difference to the game. And give feedback of how the players can improve their own performance in each game. And also give have a better planning of how the match should be played overall by the whole team regardless of the toss decision.

APPROACH AND DESIGN

The below figure explains the approach we have taken into building the predictive model using machine learning algorithms.

SYSTEM REQUIREMENT SPECIFICATION

A System Requirement Specification (SRS) is basically an organization's understanding of a customer or potential client's system requirements and dependencies at a particular point prior to any actual design or development work. The information gathered during the analysis is translated into a document that defines a set of requirements. It gives the brief description of the services that the system should provide and also the constraints under which, the system should operate. Generally, SRS is a document that completely describes what the proposed software should do without describing how the software will do it.

It's a two-way insurance policy that assures that both the client and the organization understand the other's requirements from that perspective at a given point in time. SRS document itself states in precise and explicit language those functions and capabilities a software system (i.e., a software application, an ecommerce website and so on) must provide, as well as states any required constraints by which the system must abide. SRS also functions as a blueprint for completing a project with as little cost growth as possible. SRS is often referred to as the "parent" document because all subsequent project management documents, such as design specifications, statements of work, software architecture specifications, testing and validation plans, and documentation plans, are related to it.

Requirement is a condition or capability to which the system must conform. Requirement Management is a systematic approach towards eliciting, organizing and documenting the requirements of the system clearly along with the applicable attributes. The elusive difficulties of requirements are not always obvious and can come from any number of sources.

Functional Requirements

Functional Requirement defines a function of a software system and how the system must behave when presented with specific inputs or conditions. These may include calculations, data manipulation and processing and other specific functionality.

Following are the functional requirements on the system:

1. The whole process can be handled at minimal human interaction with android and web both.
2. The application automatically receives the captured data from server.
3. The user can call emergency, map location and ECG graph on demand
4. The system gives a warning message.

Non Functional Requirement

Non-functional requirements are the requirements which are not directly concerned with the specific function delivered by the system. They specify the criteria that can be used to judge the operation of a system rather than specific behaviours. They may relate to emergent system properties such as reliability, response time and store occupancy. Non-functional requirements arise through the user needs, because of budget constraints, organizational policies, the need for interoperability with other software and hardware systems or because of external factors such as :-

- Performance Requirements
- Design Requirements
- Security Constraints
- Basic Operational Requirements

Product Requirements

- ### Platform independency:
A progressive web app will be developed and deployed so that users with a smartphone or a computer can access the voting site to cast their vote.

- ### Easy to use:
The progressive web app provides an interface which is easy to use and eliminates the need for the voter to go to a voting booth.

- ### Modularity:
The complete product is broken up into modules and well-defined interfaces are developed to explore the benefit of flexibility of the product.

- ### Robustness:
This software is being developed in such a way that the overall performance is optimized, and the user can expect the results within a limited time with utmost relevancy and correctness.

System Configuration

H/W System Configuration:

- Processor-core – III
- Speed-1.1 GHz
- RAM-256 MB (min)
- Hard Disk- 20 GB

S/W System Configuration:

- Operating System   :-- XP/7/8/8.1/10
- Coding Language   : -- Python ,HTML,CSS,Javascript.

Hardware Requirements

- Processors- Vostro IV Processor
- Speed- 3.00 GHZ
- RAM- 8 GB
- Storage- 20 GB

Software Requirements

- Operating system- Linux Mint 21
- IDE used- Visual Studio Code

SYSTEM DESIGN

Design is a meaningful engineering representation of something that is to be built. It is the most crucial phase in the developments of a system. Software design is a process through which the requirements are translated into a representation of software. Design is a place where design is fostered in software Engineering. Based on the user requirements and the detailed analysis of the existing system, the new system must be designed. This is the phase of system designing. Design is the perfect way to accurately translate a customer's requirement in the finished software product. Design creates a representation or model, provides details about software data structure, architecture, interfaces and components that are necessary to implement a system. The logical system design arrived at as a result of systems analysis is converted into physical system design.

System development methodology

System development method is a process through which a product will get completed or a product gets rid from any problem. Software development process is described as a number of phases, procedures and steps that gives the complete software. It follows series of steps which is used for product progress. The development method followed in this project is waterfall model.

Model phases

The waterfall model is a successive programming improvement process, in which advance is seen as streaming relentlessly downwards (like a waterfall) through the periods of Requirement start, Analysis, Design, Implementation, Testing and upkeep.

Prerequisite Analysis:

This stage is worried about gathering of necessity of the framework. This procedure includes producing record and necessity survey.

Framework Design:

Keeping the prerequisites at the top of the priority list the framework details are made an interpretation of into a product representation. In this stage the fashioner underlines on calculation, information structure, programming design and so on.

Coding:

In this stage developer begins his coding with a specific end goal to give a full portray of item. At the end of the day framework particulars are just changed over into machine coherent register code.

Usage:

The execution stage includes the genuine coding or programming of the product. The yield of this stage is regularly the library, executables, client manuals and extra programming documentation.

Testing:

In this stage all projects (models) are coordinated and tried to guarantee that the complete framework meets the product prerequisites. The testing is worried with check and approval.

Support:

The upkeep stage is the longest stage in which the product is upgraded to satisfy the changing client need, adjust to suit change in the outside environment, right mistakes and oversights beforehand undetected in the testing stage, improve the proficiency of the product.

System Details

Use case diagram
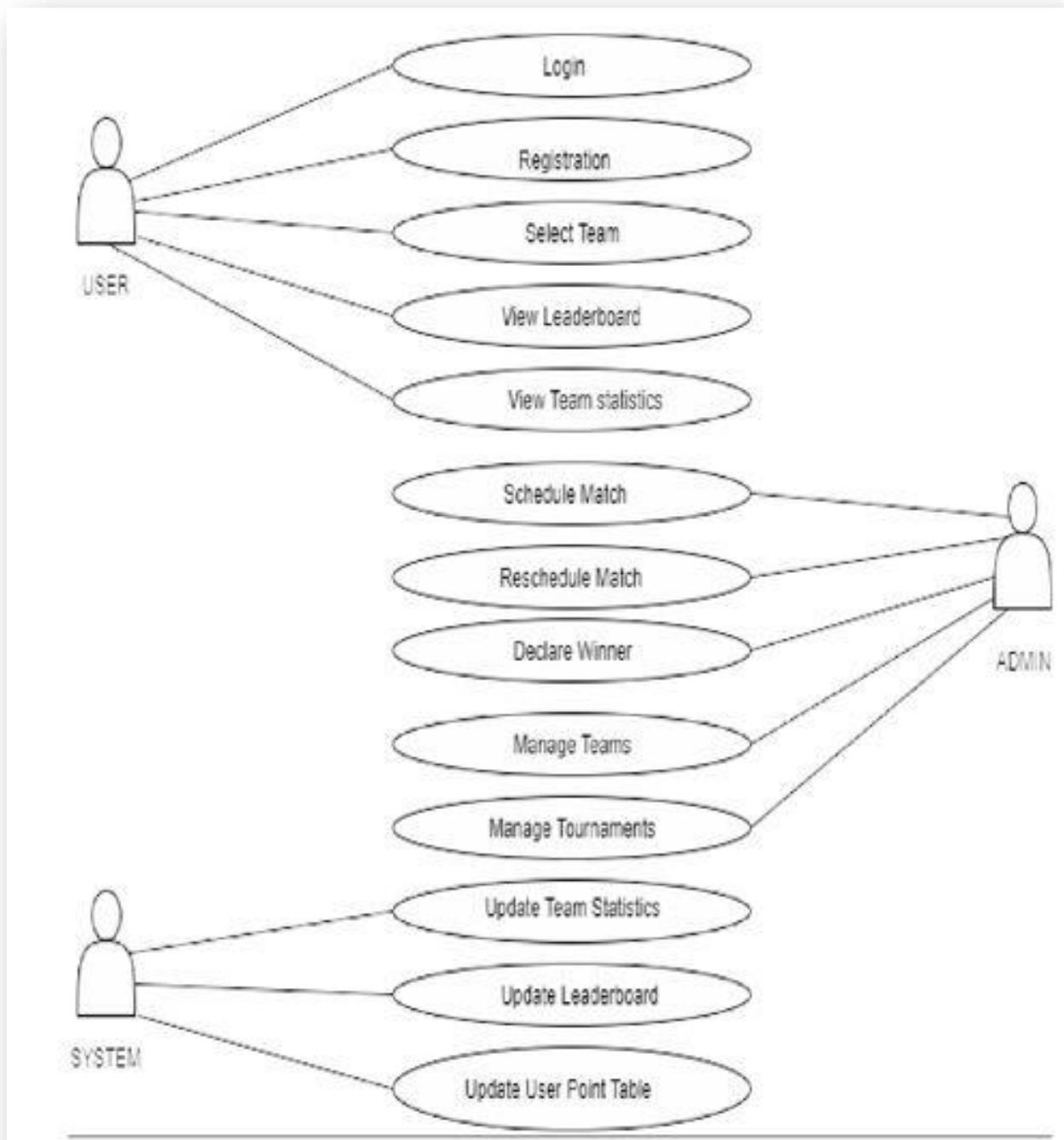
ER Diagram
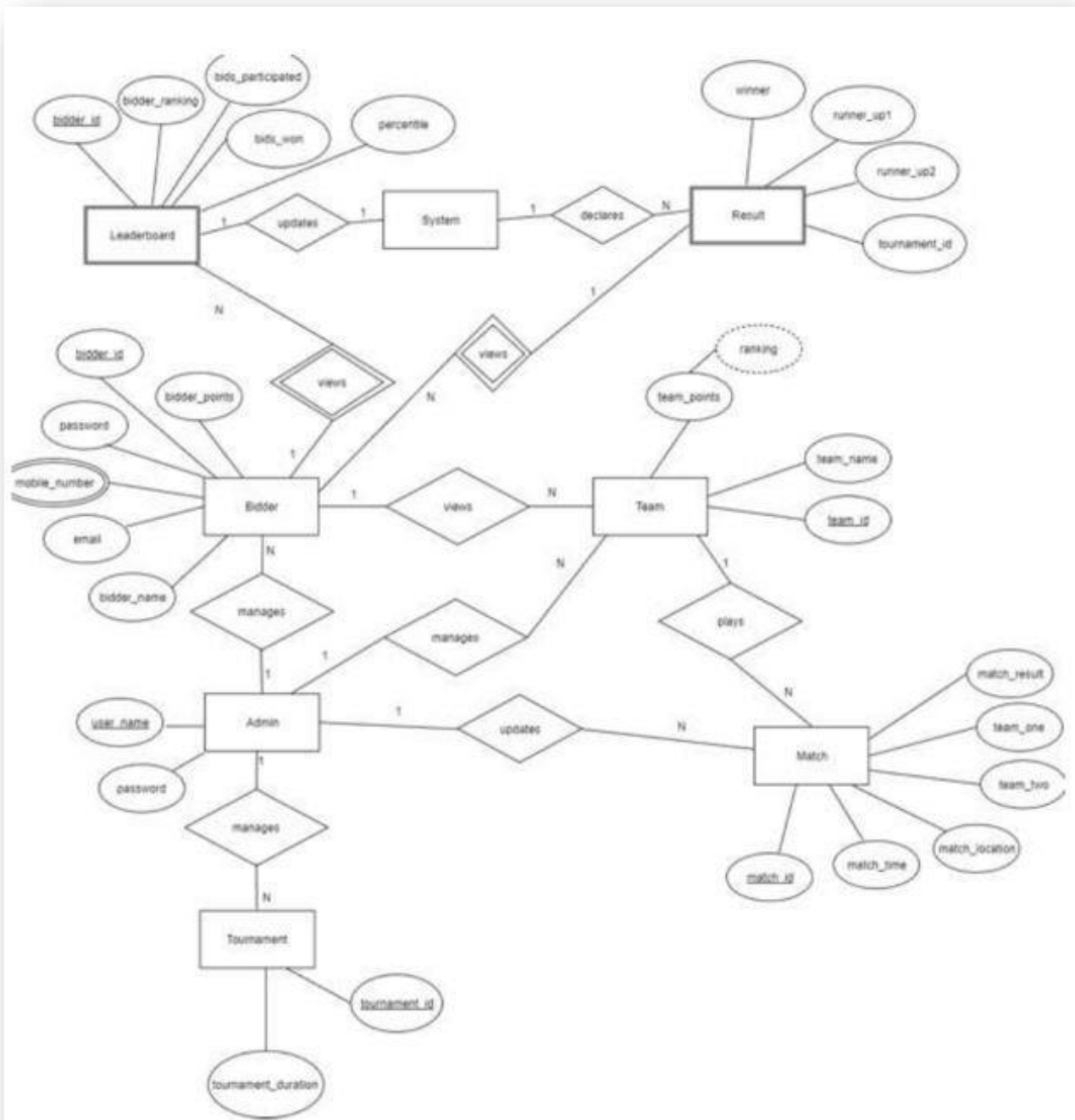
System Architecture
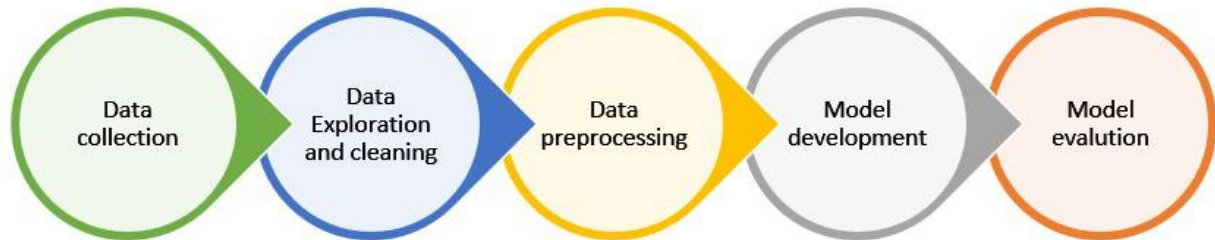
Fig. Proposed system architecture

Sequence Diagram:

Activity Diagram :

Activity Diagram For Cricket Match Outcome Prediction

User	System	Machine Learning Model
	Collect Dataset	
	Feature Extraction	
	Split Dataset	
	Testing Data	Training Data
		Train Model
Prediction Result		Test Model
New Data		Test On New Data
Prediction		
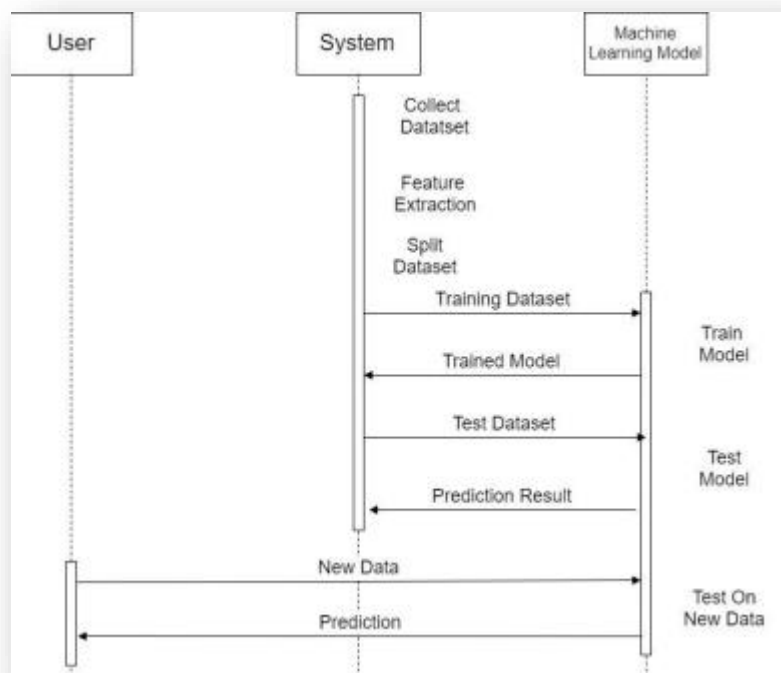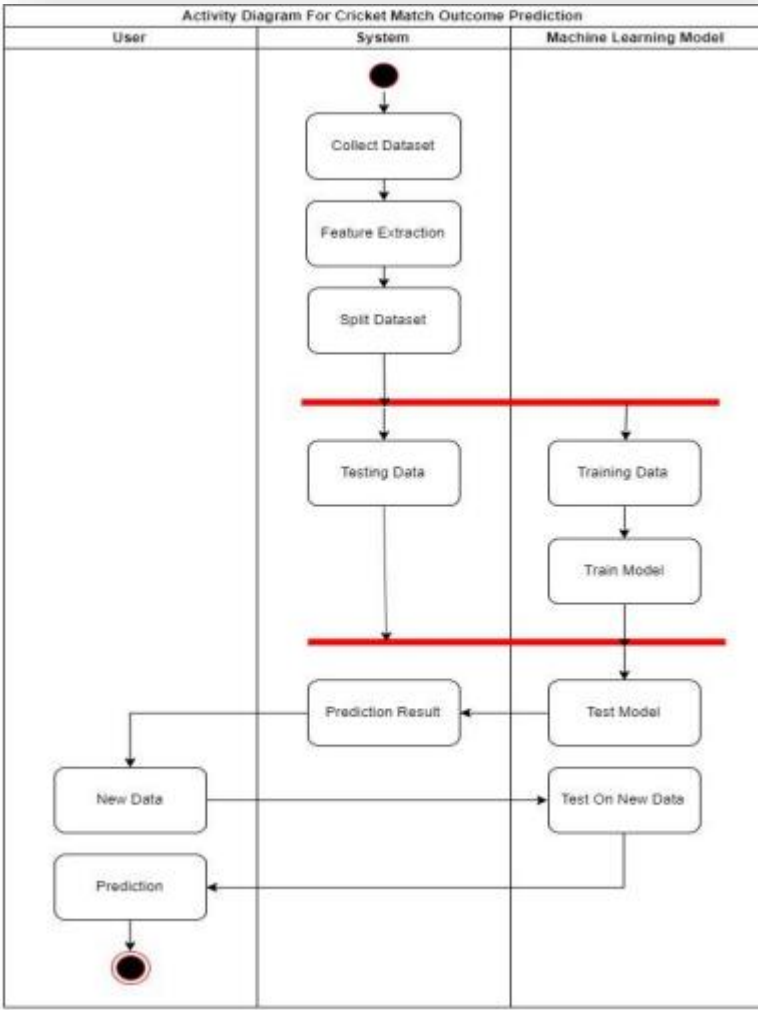
Implementation

IPL Score Predictor:

IPL Score Predictor Result:

T-20 Score Predictor Page:

One Day Score Predictor Page:
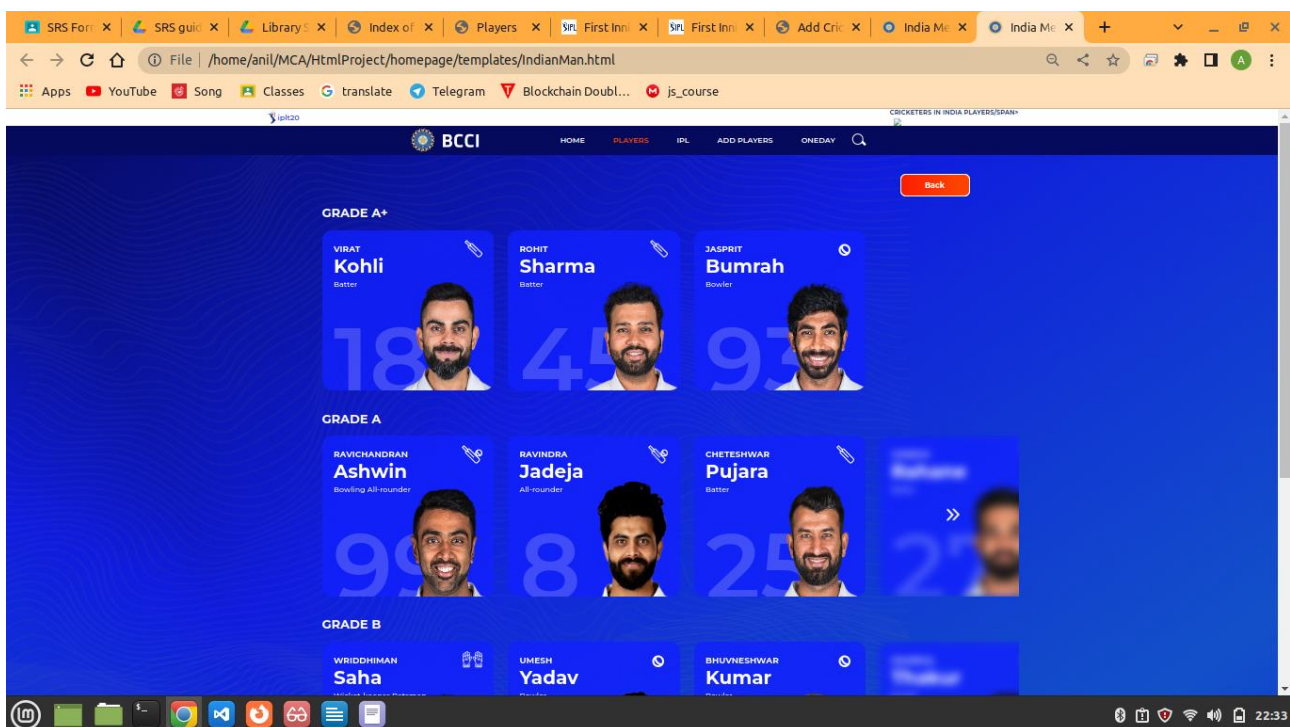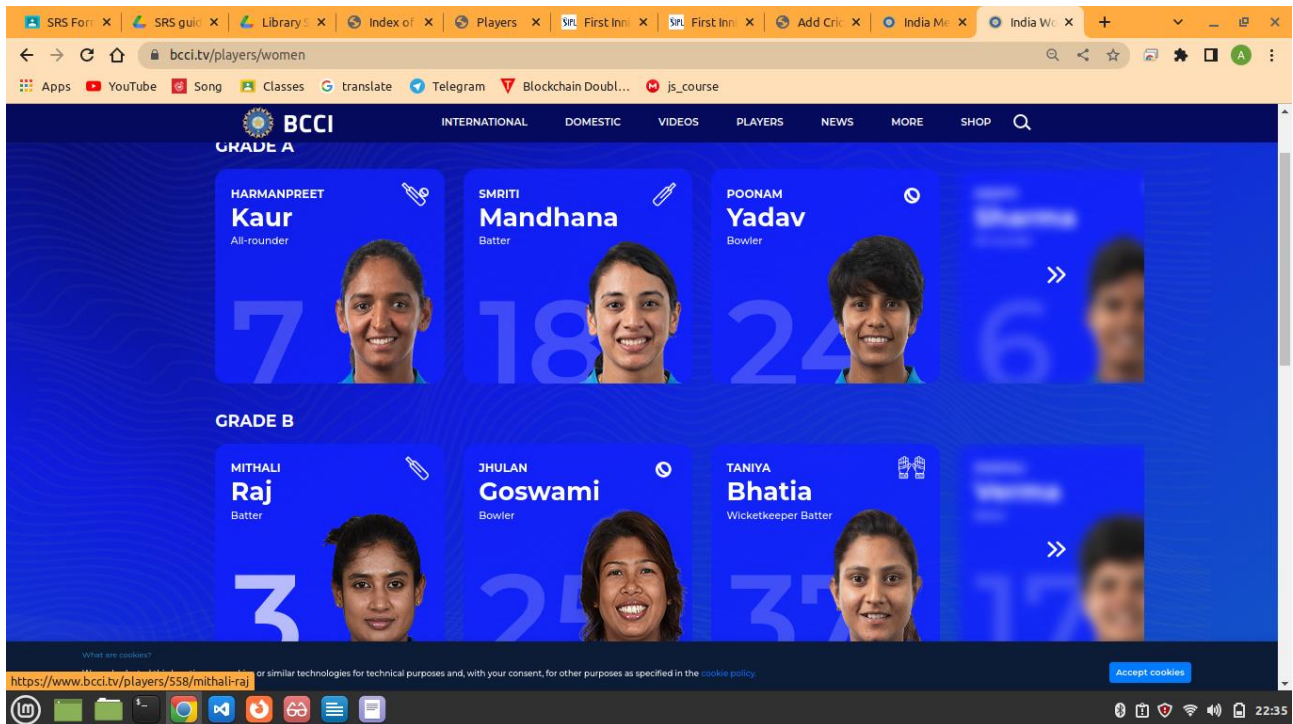
All Teams Cricketer's Players List Pages
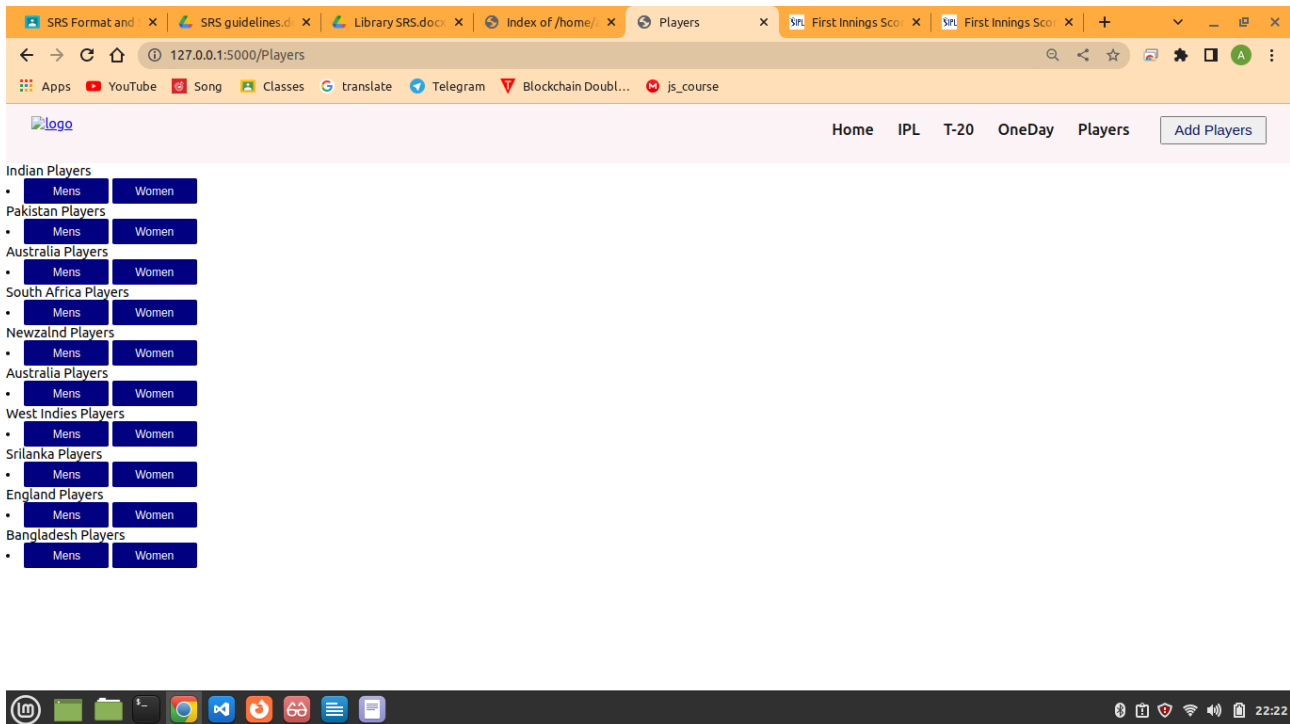
Click Button Indian Man's Players Details:

Click Button Indian Woman's Players Details:

FUTURE SCOPE AND CONCLUSION

Selection of the best team for a cricket match plays a significant role for the team's victory. The main goal of this paper is to analyse the IPL cricket data and predict the players' performance. Here, three classification algorithms are used and compared to find the best accurate algorithm. The implementation tools used are Anaconda navigator and Jupyter. Random Forest is observed to be the best accurate classifier with 89.15% to predict the best player performance. This knowledge will be used in future to predict the winning teams for the next series IPL matches. Hence using this prediction, the best team can be formed. This project opens scope for future work in the field of cricket and predicting other important things like best team of players, best venue, best city, best fielding decision to win a match.

REFERENCES

➢ T. A. Severini, Analytic methods in sports: Using mathematics and statistics to understand data from baseball, football, basketball, and other sports.

➢ Chapman and Hall/CRC, 2014.8. H. Ghasemzadeh and R. Jafari, "Coordination analysis of human movements with body sensor networks:

➢ A signal processing model to evaluate baseball swings," IEEE Sensors Journal, vol. 11, no. 3, pp. 603–610, 2010

➢ Lern-Vern Learning Platform.