

**Player Performance Prediction using Machine Learning Algorithm Case Base
Cricket in Kenya.**

Bhojani, Mirav Hirji
Student ID: 136446

**An Information System Project II proposal submitted to the School of Computing and
Engineering Sciences in partial fulfilment of the requirements for the award of the
Bachelor's Degree in Business Information Technology of
Strathmore University**

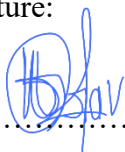
School of Computing and Engineering Sciences
Strathmore University
Nairobi, Kenya.

July 2023

Declaration

I declare that this work has not been previously submitted and approved for the award of a Bachelor's degree by this or any other University. To the best of my knowledge and belief, the proposal contains no material previously published or written by another person except where due reference is made in the proposal itself.

Student's signature:



..... [Signature]

.....14/07/2023..... [Date]

Approval

The Information System Project 2 proposal of Mirav Hirji Bhojani was reviewed and approved (*for examination*) by: Mr. Kelvin Kibunja.

Supervisor's signature:

..... [Signature]

.....14/07/2023..... [Date]

Abstract

Kenya in the last two decades has seen a major downfall in cricket. The main reasons for this downfall are, corruption, poor management and outdated technology. Thus, a web application is proposed to minimize the effect from these factors.

The proposed system will be used by Cricket Kenya to score all matches being played locally so as to maintain centralized data. After the end of any season a leaderboard will be generated which will give an invite to players to give trials to be selected in the national team. From there a 17-player squad will be announced. Then for every game a player performance prediction will be done which will then be used to recommend the best possible eleven for the team. If the combinations are chosen well the outcome of the game will benefit Kenya.

The proposed system will be developed using a machine learning algorithm, specifically a classification model that will classify the players performance prediction to either high, poor or average and thus those with high performances will be selected in the playing eleven using random forest and evolutionary prototyping methodology with the help of Object-Oriented Analysis and Design paradigm.

In conclusion, if the system is built as planned, then the intended problem of corruption will be resolved such that players not appearing on the leaderboards will not be allowed to play for the national team. Moreover, the poor team management and outdated technology and intelligence issue will be resolved using the prediction tool hence the system will overall benefit the game of cricket in Kenya.

Keywords: Cricket Kenya, 17-player squad, best possible eleven and evolutionary prototyping.

Table of Contents

Declaration.....	ii
Approval	ii
Abstract.....	iii
List of Figures	vii
List of Tables	viii
List of Abbreviations	ix
Acknowledgements.....	x
Chapter 1: Introduction	1
1.1 Background	1
1.2 Problem Statement	2
1.3 Aim.....	3
1.4 Specific Objectives.....	3
1.5 Research Questions	3
1.6 Justification	4
1.7 Scope	4
1.8 Delimitations and Limitations	5
1.8.1 Delimitation	5
1.8.2 Limitation.....	5
Chapter 2: Literature Review.....	6
2.1 Introduction.....	6
2.2 Current player selection process for the starting eleven	6
2.2.1 Challenges faced by using the current player selection process	7
2.3 Related Works.....	7
2.3.1 Data Mining System for Selecting a Winning Cricket Team	7
2.3.2 Player's Performance Prediction in ODI Cricket Using Machine Learning Algorithms	9

2.3.3 Machine learning-based Selection of Optimal sports Team based on the Players Performance	10
2.4 Gaps in existing systems	11
2.5 Conceptual Framework	11
Chapter 3: Methodology	13
3.1 Introduction	13
3.2 Methodology	13
3.2.1 Requirements	14
3.2.2 Quick Design	14
3.2.3 Build Prototype	16
3.2.4 User Evaluation	16
3.2.5 Refine Prototype	16
3.2.6 Implement and Maintain	17
3.3 Deliverables	17
3.3.1 Proposal Document	18
3.3.2 Cricket Scoring System	18
3.3.3 Prediction Model	18
3.4 Tools and Techniques	18
3.4.1 Language	18
3.4.2 Code Editor and IDE	18
3.4.3 Output	19
3.4.4 Framework	19
3.4.5 Database	19
3.4.6 Backup	19
3.4.8 API	19
3.4.9 Diagram Tools	19
3.4.10 SciKit – Learn	20

References.....	21
Appendix.....	25
Appendix 1.....	25
Appendix 2.....	26

List of Figures

Figure 2.3. 1: Features to predict batsmen's performance	8
Figure 2.3. 2: Features to predict all player's performance	8
Figure 2.3. 3: Prediction of all player's performance	9
Figure 2.3. 4: Display of the system	10
 Figure 2.5. 1: Conceptual Framework	 12
 Figure 3. 1: Prototyping Methodology.....	 13
 Appendix 1: Gantt Chart.....	 25

List of Tables

Table 1: Tests to be used.....	17
Table 2: Student Details.....	26
Table 3: Marking Guide.....	26

List of Abbreviations

API – Application Programming Interface

ERD – Entity Relationship Diagram

ICC – International Cricket Council

IDE – Integrated Development Environment

ODI – One Day Internationals

OOAD – Object Oriented Analysis and Design

T20 – TwentyTwenty Format

T20I- TwentyTwenty International

Acknowledgements

As a student, I am overwhelmed with the help and support that I received in the process of creating this proposal, especially from my supervisor Mr. Kelvin Kibunja as well as the unit lecturer Mr. Titus Kipkosgei who upon approach at any given time have helped us with appropriate response that enabled me to finish my proposal document well before the deadline. Moreover, I would like to thank Strathmore University and the School of Computing and Engineering Sciences which allowed me to express the skills and knowledge to come up with a system that could benefit the society.

Lastly, there was no viable way of completing the proposal document without the grace of God who gave me very supportive parents and guided me wherever needed despite running under busy schedules and gave me ideas to make my project unique. Finally, wanted to sincerely thank any other supporter that has unintentionally been left out in the acknowledgment.

Chapter 1: Introduction

1.1 Background

Kenya, a cricket playing country that had all the eyes of the best cricket playing countries. In the 1980's Kenya had seen growth in the game, such that the ICC (International Cricket Council) had so much faith in the country that they chose Kenya to host a World cup trophy in 1996 where Kenya managed to turn out as runners up (Mehta, 2020). Also, a World Cup knockout tournament in 2000's which is a big thing in cricket as the ICC wanted Kenya to gain for attraction and funds to develop the game in the country (Fox, 2017).

Moreover, in 2003 which is Kenya's dream year till date where Kenya managed to qualify for the final events (Banerjee, 2018). In this process Kenya managed to beat the strongest teams of that time, namely the West Indies (which shocked the whole world), India and Sri-Lanka. However, after 2003 is when Kenya started to see the decline that it lost quite a few consecutive games till 2011 where ICC then had to snatch away the ODI status of the country (Wigmore, 2014) which made Kenya to pass through qualifiers which till date they have not even been able to qualify to get a spot in the world cup (Mehta, 2020).

Currently, there are many players in the country who want to play Cricket as a profession but due to less scope, corruption, poor management, and outdated technology in the country for cricket the players opt to choose for other occupations for better pay and better opportunities and play cricket as a hobby on the side (Dr. Doreen Odhiambo, 2020).

From the above information it suggests that Cricket in Kenya is almost dead, however there is a chance of reviving it with the help of a proposed system. The system will be able to provide up-to-date and relevant information that will help fight corruption, poor management and more so use the technology to analyse players well and make informed and reliable decisions to improve the outcome of the matches played and also when selecting the national team.

1.2 Problem Statement

The proposed system will try its best to deal with the problems identified in order to revive cricket in Kenya. The web application will do scoring of all domestic matches being played in Kenya and store and maintain all the data (Banerjee, 2018). The scoring will be done on ball-to-ball basis and will update the score after every ball. For the batsman it scores which side of the ground they are getting out to, and for bowlers the side they are hit the most. Then after the season is over batsman scores per match are tallied up and total runs added up and a ranking leader board will be generated with the highest run scorer as number 1 and the rest follow. The same for the bowlers, the wickets per match will be accumulative and will be added at the end of the season with the highest wicket taker being ranked as first and the order follows. This system will also allow cricket Kenya to have centralized data that is stored and managed by them (Gavhane, 2021).

Once the leader boards are generated, the top seven players on the list of both batsman including the top three wicket keepers and bowlers will be sent an invite for trials. This will shortlist a 17-man squad for any upcoming series or match. Before any matches these 17 players will be keyed into the system to predict the performance of their upcoming matches and the prediction will be in three verdicts that is good, poor and average (Kumar, 2021). This prediction will be based on the players' previous data on certain grounds, opponent team, average and many other related factors. This data will be trained and this will check the previous performances for the last five matches and will check the trend and give an outcome. The players with the good performance prediction will be selected in the starting eleven of the matches. The players in the starting eleven will be predicted to perform well and if the prediction is accurate then the outcome of the match will be favourable to cricket Kenya (Kalpdrum Passi, 2018). This will lead to development of the game and of players as well and deal with the issue of corruption and management since the players in the leader boards will only be allowed to play for the national team.

However, the game has different formats and categories and due to time constraints only one of the formats will be able to be focused on and that is the 5 over format for the main men's team and using a web application. This is because 5 overs are the shortest format of the game and it still takes 1 and a half hour which already is a lot of time and thus the other formats are even longer and have more complexities with the rules. 5 overs format will be 5 overs a side and thus is preferred by most players (Mohanta, 2020).

1.3 Aim

To design and develop a web application that predicts player's performance to enhance the selection of starting eleven and deal with the corruption issues to revive cricket in Kenya.

1.4 Specific Objectives

- i. To examine and understand the current player selection process.
- ii. To review the challenges faced by the existing selection process.
- iii. Reviewing machine learning algorithms used in player selection.
- iv. To develop a classification model to predict player performances for future matches.
- v. To test the proposed system using unit and integration testing.

1.5 Research Questions

- i. What is the player selection process?
- ii. What are the challenges faced by using the existing process?
- iii. What are the various machine learning algorithms that can be used to select players?
- iv. How effective will the option of developing the model using a classification algorithm?
If not, which other algorithm could be used to meet the system requirements?
- v. What type of testing will be used to check on the validity of the system to be developed?

1.6 Justification

The current selection process of the starting eleven is based on a few assumptions and gut feelings and how close a player is to the coach (Mehta, 2020). Also, the issues of corruption and use of outdated technology are still in place and also the collection of data is through external apps like CricHQ and CricHeroes (Jain, 2023) which leads to inconsistency in the data collected (CricHQ, 2018) (Desai, 2022).

In order to bridge the issues brought up by the current processes the proposed system that deals with the issues of corruption, outdated technology and centralization of the data storage and all this will be achieved by developing an application that does scoring for all the matches being played in Kenya which at the end updates a leader board for top performers who are only the ones allowed to play for the national team unlike the players who initially were in the team and with poor recent performance. From the leader boards a 17-man squad is generated including 7 bowlers, 7 batsman and 3 wicketkeepers (Kendall, 2023). The starting eleven will then be selected using a prediction model that will predict the players performance for future endeavours based on previous match performances. The players who are predicted with a verdict of either good performance or average performance. This will both enhance the process of selection of the starting eleven and return a favourable outcome of the match which in the long run will benefit the game in the country and help in development of players skills (Aneem-Al-Ahsan Rupai, 2020) and also be an opportunity for other players with natural talent be able to choose cricket as a profession unlike currently since there is no scope, people are fearing to take cricket as an option for their career paths (Okinyo, 2020).

1.7 Scope

As there are many formats of cricket the proposed project will focus on the 5 over side format which is one of the most popular formats of the game for street cricket and very interesting as well (Mohanta, 2020). Together with only matches being played locally that is in Kenya only. The reason for going with the 5-over format is because the data for this format is easily available through CricApi.

1.8 Delimitations and Limitations

1.8.1 Delimitation

Although Kenya does not have the status to play the fifty-fifty format and the Test format, they still play these formats under the supervision of ICC however no points are given to Kenya for playing these matches (Thiraviam, 2023). This is therefore a delimitation as there is then no point on focusing what the country does not play. This could be a project for the future when cricket in Kenya reaches its highest again and get the Test format status.

1.8.2 Limitation

The limitations of the project will be that the project will only be able to base the country to Kenya since Kenya has not played many international matches and the data for the international played are not available and thus the project will be limited to Kenyan boundaries.

Chapter 2: Literature Review

2.1 Introduction

This chapter explains all the current player selection processes and the challenges faced by the current problems and reviewing the existing systems to identify and learn the methods used by them, list out the gaps in these existing projects and what can be done to reduce the bridge between these gaps identified and finally the conceptual framework of the proposed system to suggest how the system will work and the roles of the major users of the system.

2.2 Current player selection process for the starting eleven

With more than 100 countries playing professional cricket (Fast Advices, 2019), all the countries have to think well through the selection process for the starting eleven for the matches so as to get the best outcome for the match. Even after a lot of brainstorming, the reason why most of teams fail to choose the best combination for their starting eleven is due to lack of intelligence. Mostly, the selection is done by the captain, the coach and the selectors and sometimes the decision made on specific players does not come out as expected and because of this the outcomes of the match sometimes turns out to be unfavourable (Your Wise Information, 2022).

Players in the starting eleven are selected based on many factors like past performances, home ground or away, experience in the match, venue specified performance, against specific team performance and the form of player (Shah, 2020).

Usually, the selection is based on experience that is the most experienced players are given priority in their respective roles and after them follow those who are currently in form. However, usually the captain of the team has the final call on the selection after listening to the advice of certain players from the coaches and the selection panel after doing all the analytics (Jayanth, 2018). Sometimes the analytics could be wrong and thus the combination comes out to be unfit for the situation and thus the match turns out to be unfavorable for the team.

The same is the case with the selection process in Kenya. The players with most experiences are given priority followed by the players who have been performing well in the domestic tournaments who are anticipated to play well in the international matches.

2.2.1 Challenges faced by using the current player selection process

From time-to-time existing processes need to be reviewed in order to keep up with the rest of the world and their advanced processes. Kenya over the years has been reviewing their selection process since 2010 and still is struggling to get the right process (ESPN, 2010). Again in 2021 they wanted a few changes in the process, this shows that the process has been very unreliable and constantly needs to be reviewed (Ayodi, 2021).

There are many challenges faced by using the current player selection process as it is a very cumbersome and very important process there cannot be any compromises with it otherwise the outcome of the match will turn against the team. Some of the challenges are that in a pre-selected squad there almost two to three players for every slot with almost the same experience and thus it gets hard to select one player over the other and then this leads to gut feeling selection which not always is the correct decision. Moreover, the players who have a good relationship with the captain have a higher chance of getting selected in the team which is not fair to other deserving players (Frontiers in Psychology, 2021). Also, the conditions of the pitch and the other factors mentioned above also affect the choosing of the players, these conditions are evaluated then a decision is made and this decision can sometimes be based on a false evaluation and this leads to maybe even losing the match (Rajesh, 2018).

2.3 Related Works

2.3.1 Data Mining System for Selecting a Winning Cricket Team

This is work done by a group of 3 students from Sri Lanka who came up with an idea of developing a system that predicts players performance to help select a starting eleven team. Their focus was based on certain factors such as previous performance, height, style, average, strike rate, number of matches and number of innings played. From the above-mentioned factors, they used the data and predicted if a specific player will perform well in the upcoming matches. This was achieved by developing a regression model specifically with Random Forest regression, Decision tree regression, KNN regression and Linear regression. The final outcome turned out to be accurate using the random forest regression however one of the issues was that some of the features used did not make sense to the prediction such as batting style or bowling style which has no direct relation to performance as what matters is skills which both a left hand or a right-hand style player can accumulate through practise and training. Thus, this leads to

over fitting of data which can bias the outcome of the match or even the performance of the player (Dinithi Hasanika, 2021).

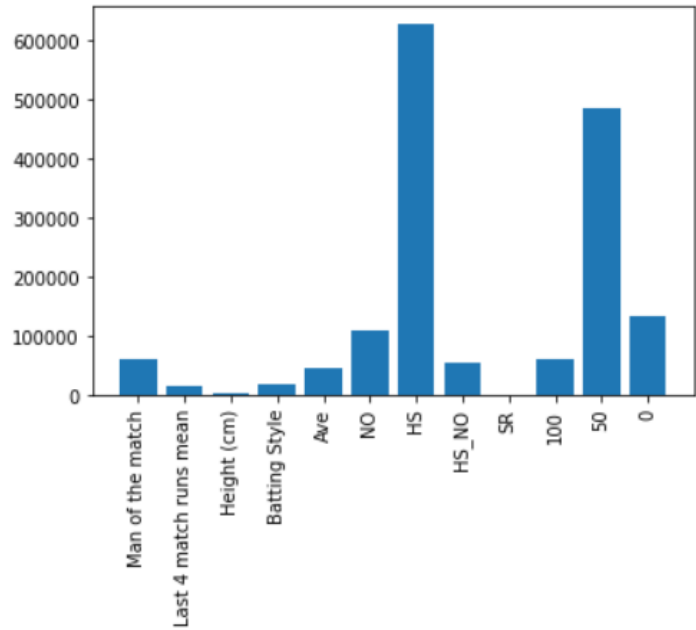


Figure 2.3. 1: Features to predict batsmen’s performance

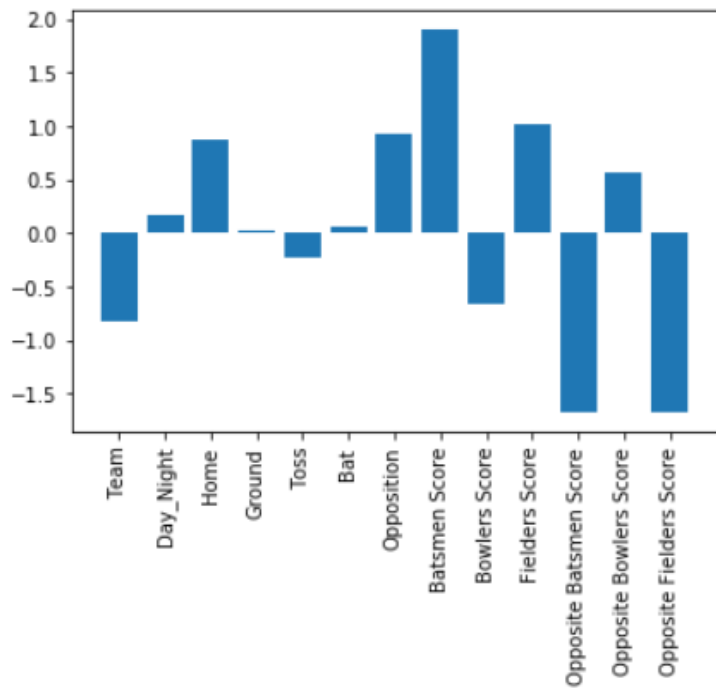


Figure 2.3. 2: Features to predict all player’s performance

2.3.2 Player's Performance Prediction in ODI Cricket Using Machine Learning Algorithms

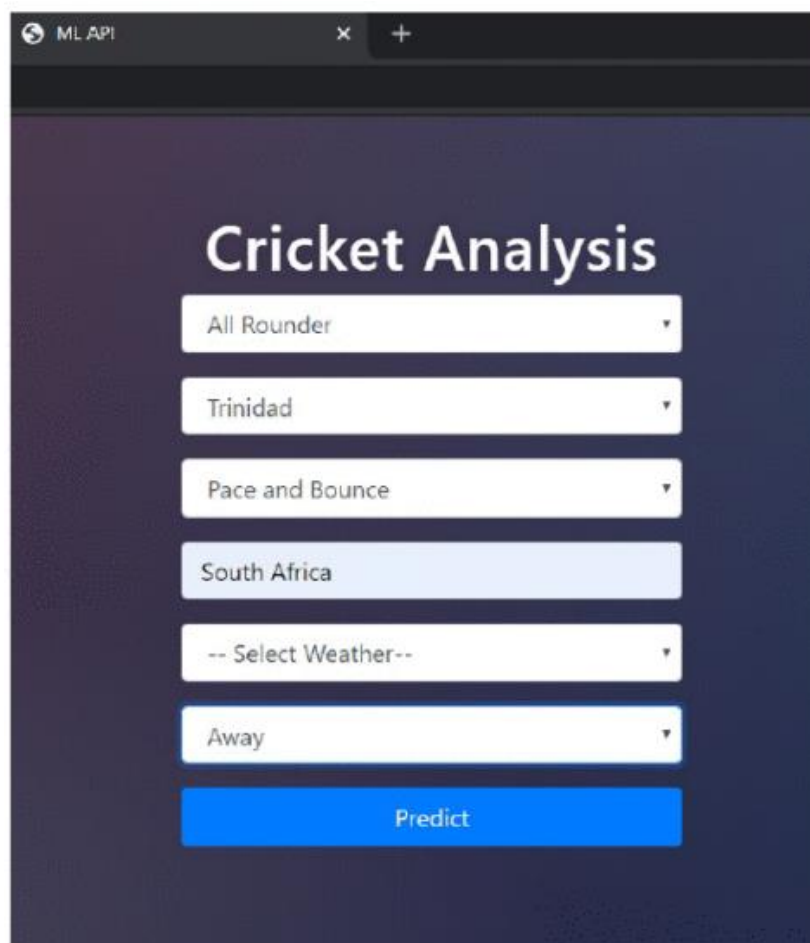
The aim of the above paper by the authors was to predict performance of a player to allow them to enable selecting a team with the best combination to help with the satisfactory outcome for the match. The algorithms used to develop this system were Linear regression and SVM with Linear and Polynomial Kernel. One of the few limitations in their project is that the features they are using to consider the prediction are not accurate and they require more accurate features such as previous few scores to know where a player is in current form or not. Despite this limitation the project has turned out to be accurate with minimum errors (Aminul Islam Anik, 2018).

Name	BF (Ball Faced)	Pos	Opposite	Ground	Pitch	Run	Linear	Accuracy
Tamim Iqbal	102	1	5	2	9	84	83.97614	99%
Shakib-Al- Hasan	63	3	5	2	9	67	52.71177	79%
Mushfiqur Rahim	52	4	5	2	9	62	41.66451	67%
Mahmudullah Riyad	23	5	5	2	9	24	15.46297	62%
Mashrafe Mortaza	5	7	5	2	9	6	5.363392	89%
Nasir Hossain	1	8	5	2	9	0	0.978821	97%

Figure 2.3. 3: Prediction of all player's performance

2.3.3 Machine learning-based Selection of Optimal sports Team based on the Players Performance

The objective of this study and project was to help in selecting the team with the perfect combination of players to win the match. The technologies and algorithms used for this project are logistic regression, SVM and random forests. The data was obtained from IPL matches and trained for Indian players. Since it was trained and tested in Indian conditions and players it is then limited to Indian boundaries thus limiting the scope of the whole project to the whole world (Monali Shetty, 2020).



The image shows a web browser window with a single tab titled 'ML API'. The page has a dark blue background and is titled 'Cricket Analysis' in large white text. Below the title, there are six input fields, each with a dropdown arrow on the right. The first field contains 'All Rounder', the second 'Trinidad', the third 'Pace and Bounce', the fourth 'South Africa' (which is highlighted with a light blue background), the fifth '-- Select Weather--', and the sixth 'Away'. At the bottom of these fields is a large blue button with the text 'Predict' in white.

Figure 2.3. 4: Display of the system

2.4 Gaps in existing systems

Majority of the existing systems relating to player performance prediction have an issue with the features used to predict. Some of these features are either unnecessary or some are missing and need to be there for the prediction. It is mostly the issue of overfitting or underfitting like for some they have features relating to the physical body such as height which technically does not have any direct effect on the game as it then biases the game to those who are tall which is not right for any prediction (Dinithi Hasanika, 2021). Also as these projects are developed and trained in one country they are then limited to benefit that country only for example Player's Performance Prediction in ODI Cricket Using Machine Learning Algorithms was a project developed in Bangladesh thus it can only benefit Bangladeshi players and the country and not the world at whole (Aminul Islam Anik, 2018) and Machine learning-based Selection of Optimal sports Team based on the Players Performance which was developed in India so these projects then limit their scope to their respective countries (Monali Shetty, 2020). Therefore, in order to have the systems working well there is need to have all the appropriate features and be tested and trained with data all over the world to support all those teams playing the sport (Fast Advices, 2019).

2.5 Conceptual Framework

The proposed system has 3 users that is the admin, the club and the spectators. The main role of the admin is to register clubs, start or end a season and view leaderboards to generate a leaderboards file that will be sent to the AI model to train the players and their previous performance and so that a suggested starting eleven can be predicted and used for the upcoming matches. The main role of the club is to register the players laying for their club and doing the scoring for the matches being played by their club in order to provide the performance data of players to the admin or Cricket Kenya Board for them to use this data to predict the starting eleven. The spectator is a minor user who can only view scores of matches being played and previous matches just to keep up to date with what is happening in the season that is being played.

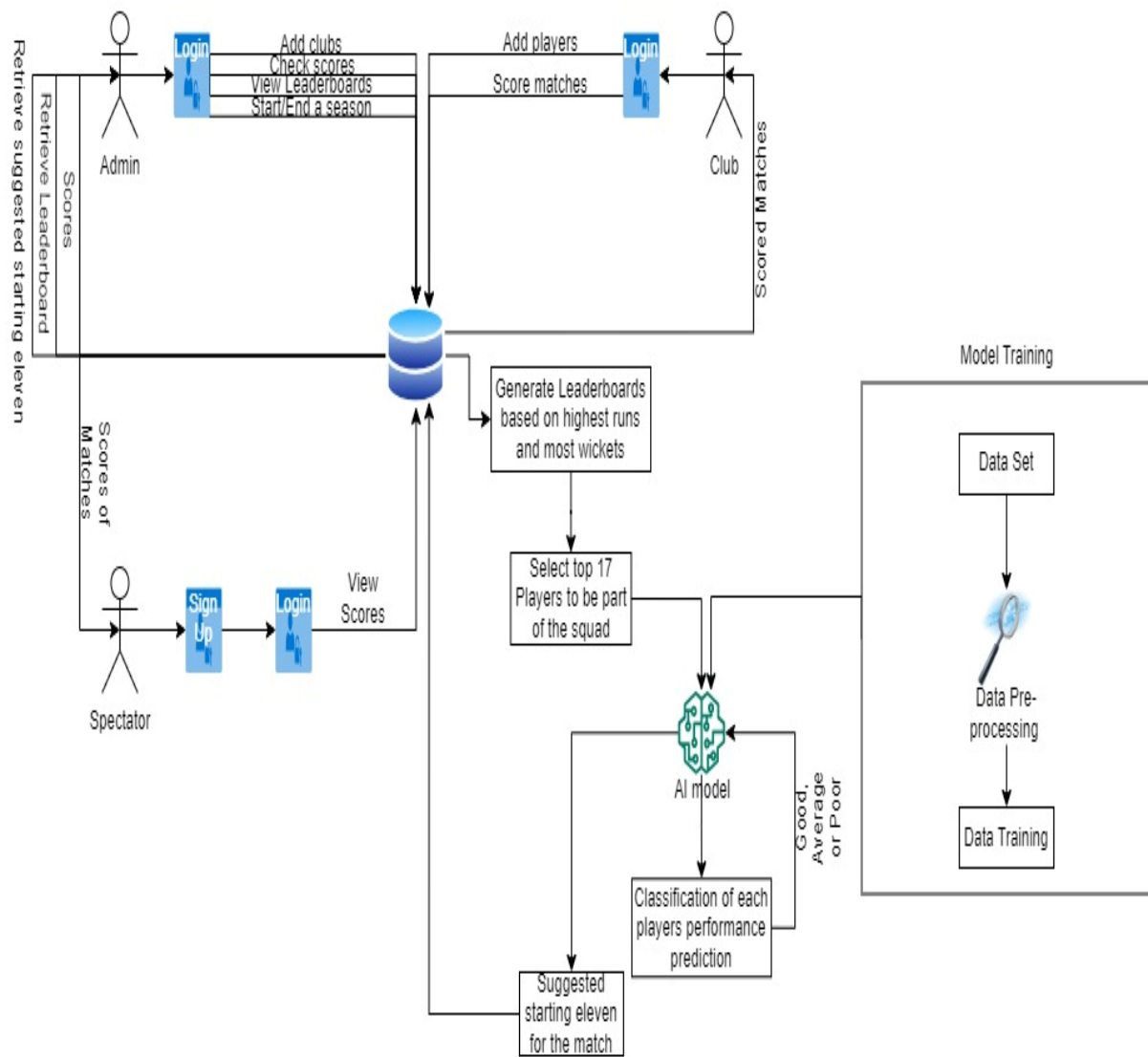


Figure 2.5. 1: Conceptual Framework

Chapter 3: Methodology

3.1 Introduction

This chapter talks about the methodology chosen for the proposed system which is prototyping explaining each phase of the methodology and the respective diagrams that are supposed to be drawn since the OOAD approach is chosen and the type of model techniques the system will use in order to develop the final product with the tools and techniques that will be used and the types of tests that will be carried out.

3.2 Methodology

The methodology that will be used to develop the system will be Prototype. The reason for choosing this methodology is that it allows to develop a mock system before actually rolling it out to the users and this mock system can cater for any changes required by the user and modify the mock system instead of coding the whole system again with the required changes (Lewis, 2023). Moreover, this methodology will also allow to identify and correct errors beforehand (Liu, 2017). Prototyping methodology turns out to be very effective since in case of any changes requires to the system then the model can be changes with the required changes unlike methodologies such as waterfall methodology (Leeron Hoory, 2022) which are less flexible to changes since they are required to finish a certain stage before moving to the other stage and then they cannot go back to a previous stage thus makes it less flexible.

Evolutionary Prototyping methodology will be specifically used since it allows to work in an iterative cycle with design and implementation to cater for the feedback from end users. Also, because it helps the reduce the headache of developing from scratch as the existing prototype can be modified to fit the requirements of the final user (Gurung*, 2020).

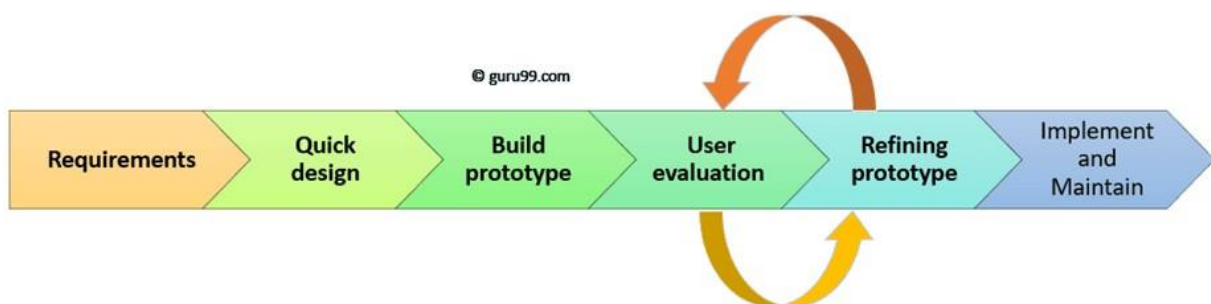


Figure 3. 1: Prototyping Methodology

3.2.1 Requirements

The methodology starts with this phase where the capabilities features and constraints of the system are identified in detail in order for the system to run smoothly. These requirements are usually gotten from following software engineering and requirements gathering process mostly by interviews as a form of data collection with ideation (Team Asana, 2022).

The data will be obtained from the scoring system and the testing will be done with data obtained from cricinfo stats provided from Kaggle ([Cricket ODI - Players Performance | Kaggle](#)).

3.2.2 Quick Design

In the second phase of the prototype methodology a basic design of the system is developed in order to facilitate in development of the prototype. This is done by drawing design diagrams that helps the developer of the system to understand the requirements well and help in fastening the development process of the prototype (Naimishsahu08, 2021). The paradigm approach chosen is Object Oriented Analysis and Design (OOAD) this is because OOAD turns out to be more suitable with prototyping methodology as mentioned above since prototyping is an iterative methodology. Moreover, this approach divides the system into use cases thus understanding the interests of specific users of the system well especially for the bank application that has users with specific roles thus understanding each user requirements well, this is easier as this approach is also data oriented. Moreover, if there are any changes or upgrades required for the system it was easy to implement (Alam, 2017). Also, OOAD uses bottom-up approach which allows developing complex systems in a simpler many easing up the development process (Hammad, 2020). Below are the analysis diagrams that will be drawn in order to help in facilitate development.

The proposed model is a machine learning models that will use supervised learning and form a classification model that learns the patterns and relationships from the input of data provided and the same will be used for the proposed system with the specific technique of decision trees and random forest this is because these techniques allow handling non linear relationships and missing data moreover allowing interpretability of the system and this then allows informed and reliable decisions to be made and accurate predictions of the players performance (MELTZER, 2021).

3.2.2.1 Use Case Diagram

The reason for drawing this diagram is to help understand the functional requirements of the system and specific users of the system that is the admin, the club, and the spectator, and developing a system keeping the interests of users in mind (El-Attar, 2019).

3.2.2.2 Sequence Diagram

The importance of a sequence diagram is to show the interactions between the users and the system by the order in which they occur. For the proposed system, the sequence diagram helped to know the processes, and which one occurs first. For example, the spectator logs in so the process starts with them entering the credentials, the credentials being sent to the database to verify if they are matching sends a message of successful login and gives access to the index page. If the credentials do not match send a message saying incorrect email or password (Creately, 2022).

3.2.2.3 System Sequence Diagram

This is a graphical representation of the system that shows how the components of the system interact with each other and generally explaining how the overall system works. They are used hand in hand with Activity diagrams (Maurya, 2020).

3.2.2.4 Entity Relationship Diagram

An ERD is drawn to help the developer understand how the data is stored in the database with the specific details related to that data hence this is the basis on which the generic database is developed (Jr., 2017).

3.2.2.5 Class Diagram

This diagram shows the visuals of the system and the relationship of data stored in the database as the relationships explained using the ERD are not well understood thus a class diagram brings out the clear relationships of data collected. This diagram allowed the developer to understand the exact responsibilities of each user of the system (Tutorialspoint , 2022).

3.2.2.6 Activity Diagram

This is a diagram that is used to describe the detailed flow from stage to another and how each activity is carried out at each stage to dynamically explain the major aspects of the system (Visual Paradigm, 2019).

3.2.3 Build Prototype

In this stage a simple system is built with the information gathered from the quick design and after understanding the requirements of the system. This system will act as a foreseen final system and sent to the users of the system for evaluation (T, 2020). Below are the tools and techniques used to develop the system.

3.2.4 User Evaluation

In this phase the users of the system assess the prototype created and provide necessary feedback after having a look at the system and noting down what needs to be added and what needs to be removed or what need to be changed before it gets too late. Moreover, they also let the developer know if there are any errors in the system (ENGINEERING DEPARTMENT | PACIFIC RESEARCH LABS, 2022).

3.2.5 Refine Prototype

Based on the feedback received from the user the prototype is modified till it fits the feedback given and again sent back to the user for evaluation and this process keeps on repeating itself until the user is satisfied completely (Pal, 2021).

3.2.6 Implement and Maintain

This is the final phase of the prototype methodology where once the system is satisfying to the user the final developments are done and tested for final errors and then deploys the website to the user's server and maintains it with any updates required in future (T, 2020). The types of tests carried out to test the final system will be the ones stated below.

Table 1: Tests to be used

Type of test	Reason.	Tool used.
Black box testing.	To check if the system is outputting correct data by data input (Rai, 2019).	Microsoft Word – to help come up with test cases. UFT Automating tool.
Alpha testing.	To validate the system before rolling it out to the market or before submitting it (Kritika, 2022).	Google. VSCode.
Regression testing.	To check if any additions in the code have affected the output of other code (Kritika, 2022).	Micro Focus Unified Functional testing.
Testing of code.	To check if the code written out is accurate such as there are no errors and not many white lines wasted (Kritika, 2022).	WebdriverIO
Automated testing	To help fully check the code and the system with minimum effort (PortSwigger LTD, 2022).	Burp Suite Application Security Testing Software

3.3 Deliverables

These refer to the specific end results that are measured and submitted or presented to the supervisors or the seniors in order to steps of completing the project and checking the progress of the project (Indeed for employers, 2020).

3.3.1 Proposal Document

This is a document that outlines everything in detail about what the whole system will be anticipating to do and how it will be done. This needs to be submitted as a deliverable since it needs to be assessed by supervisors so that they can give a go ahead on developing the actual system.

3.3.2 Cricket Scoring System

This will also be a deliverable since this system will be required for the prediction model to work as it is from here that the data for the top performers will be obtained from.

3.3.3 Prediction Model

Finally, this deliverable will be required to complete the project since the entire project is based on prediction model of a starting eleven of a cricket team.

3.4 Tools and Techniques

3.4.1 Language

The languages used for developing the website will be in a combination of HTML, CSS and PHP for the frontend of the website and for the backend there will be use of JavaScript and for the prediction model python will be heavily used. The combination of these languages used will return the best possible outcome of the project (Reshko, 2020).

3.4.2 Code Editor and IDE

The code editor and IDE that will be used to compile and run the code will be VSCode for the frontend and backend development of the web application while to model the prediction JupyterLab will be used. This is because JupyterLab is most suited for python code and allows smooth development (Singh, 2022)..

3.4.3 Output

For the output of the project Chrome will be used since it is fast, secure and reliable and also widely used thus will be easy to present using Google Chrome (Williams, 2021).

3.4.4 Framework

Laravel will help in better connecting with the database and to get better web performance and layouts compared to other frameworks. Also, Laravel allows reusability of code since any user can understand the flow of the code (Sharma, 2020).

3.4.5 Database

MySQL allows systems to store data, especially usernames and passwords, and other data with appropriate security and safety stored data. MySQL is also used to clean the data and is used as a data warehouse or data lake and is better for storage (Madur, 2021).

3.4.6 Backup

GitHub is the best place to back-up incomplete code since it is a secured form of storage with easy usage (Kachur, 2018).

3.4.8 API

This API is used to hold in building and testing the accuracy of reports created by the system and also to allow testing of the SMS integration (Hooda, 2021).

3.4.9 Diagram Tools

Draw.io and MS Project is used to help in drawing the design diagrams and the Gantt chart (Fergusson, 2016).

3.4.10 SciKit – Learn

This is a free machine learning library for python that allows access to many functions that ease up development of the model (Nik, 2022).

References

- ENGINEERING DEPARTMENT | PACIFIC RESEARCH LABS. (2022). Understanding the 5 Stages of Prototyping. *PACIFIC RESEARCH LABORATORIES*.
- Alam, R. (2017). Object Oriented Analysis and Design (OOAD). *Cybarlab*, 1.
- Aminul Islam Anik, S. Y. (2018). Player's Performance Prediction in ODI Cricket Using Machine Learning Algorithms. *IEEE*, 1109 -1110.
- Aneem-Al-Ahsan Rupai, M. S. (2020). Predicting Bowling Performance in Cricket from Publicly Available Data. *ACM Digital Library*, 1-6.
- Ayodi, A. (2021). Cricket players want independence in selection process. *Nation Prime Sports*, 3-5.
- Banerjee, S. (2018). Sandipan Banerjee. *Cricket Soccer*, 5.
- BBC Sports. (2005). ICC discuss Kenya and Test format. *BBC Sports*.
- Bell, D. (2004). UML's Sequence Diagram. *developers work*, 1-21.
- Brearley, M. (2009). How to select a cricket winning team. *Pitch Vision Sports technology*, 23-34.
- Carl J. Petersen, D. P. (2009). Movement patterns in cricket vary by both position and game format. *Taylor & Francis Online*, 45-52.
- Chappell, R. (2006). Kenyan Sport in the Aftermath of the Commonwealth Games, Manchester 2002. *Taylor and Francis Online*, 175-191.
- Creately. (2022). Sequence Diagram Tutorial – Complete Guide with Examples. *Creately*.
- CricHQ. (2018). Scoring - an intro. 3-4.
- Desai, A. (2022). 2,000 Matches Are Streamed Live In CricHeroes. *BW BusinessWorld*, 14-17.
- Dinithi Hasanika, D. L. (2021). Data Mining System for Selecting a Winning Cricket Team. *Projects- Department of Computer Engineering, University of Peradeniya*, 34-45.
- Dr. Doreen Odhiambo, N. J. (2020). Implications of Sports as An Academic and carrer pathway in the New Kenyan Curriculum Reform. *The Cradle of knowldge*, 90-96.
- El-Attar, M. (2019). Evaluating and empirically improving the visual syntax of use case diagrams. *Journal of System and Software*, 136-163.
- ESPN. (2010). Cricket Kenya revises selection process. *ESPN cricinfo*, 34-36.
- Espn Cricinfo. (2010). Cricket Kenya revises selection process. *ESPN*, 11-13.
- Fast Advices. (2019). How many countries have professional cricket teams? *FastAdvices-The best adviser for everything*, 7-9.
- Fergusson, K. (2016). Why draw.io Is Better Than Gliffy. *Seibert Weblog*, 2.

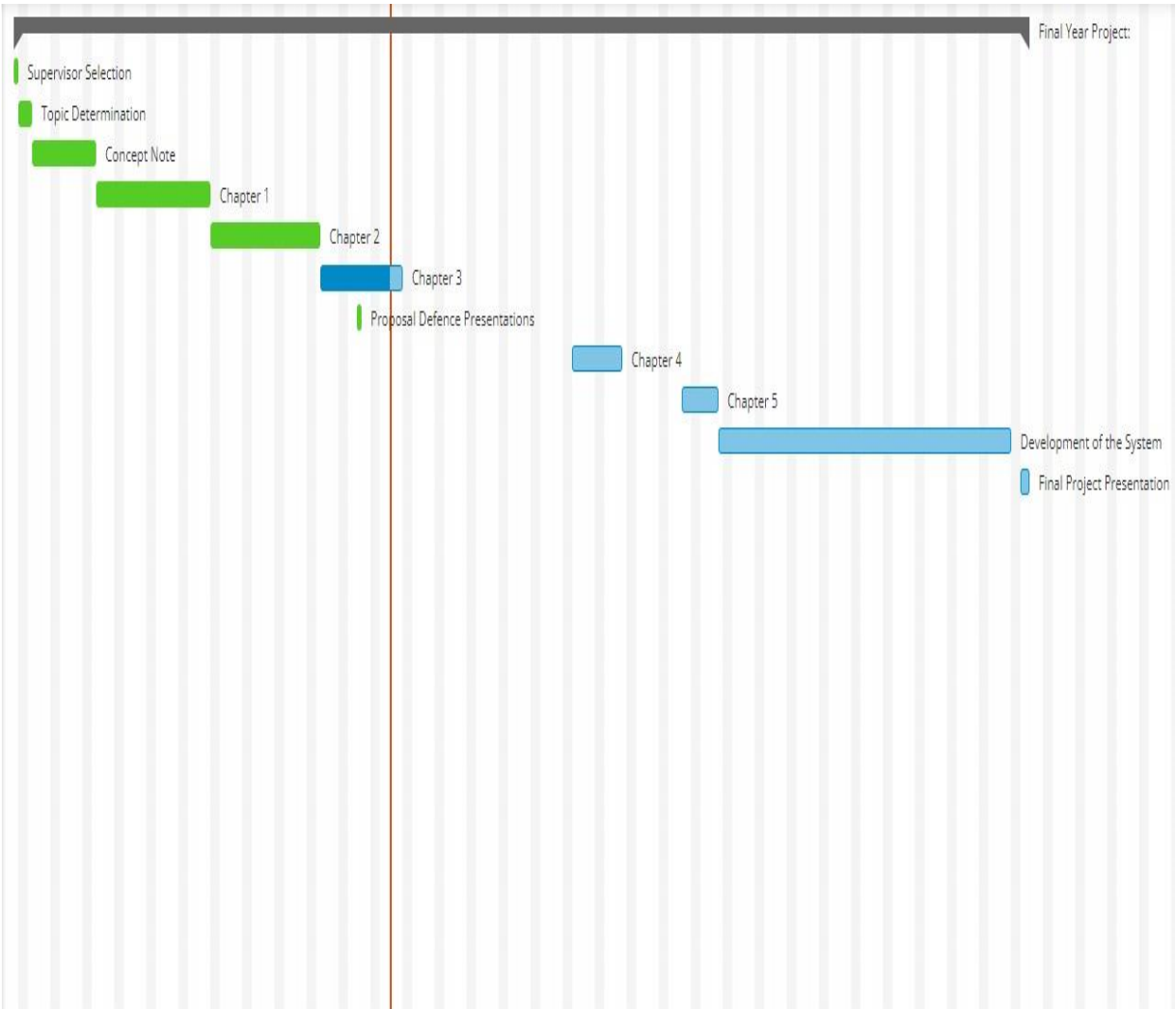
- Fox, G. (2017). Gendered issues and contested benefits. The case of Laikipia's Maasai Cricket Warriors. *HAL*, 2-4.
- Frontiers in Psychology. (2021). *Player-Character Relationship and Game Satisfaction in Narrative Game: Focus on Player Experience of Character Switch in The Last of Us Part II*, 23-28.
- Gavhane, S. (2021). Live Tournament Scoreboard and Prediction. *International Journal of Computer Applications*, 183-187.
- Gurung*, G. (2020). Software Development Life Cycle Models-A Comparative Study. *ijsrcseit*, 30-37.
- Hammad, M. (2020). Difference between Structured and Object-Oriented Analysis. *GeeksforGeeks*.
- Hooda, P. (2021). Introduction to Postman for API Development. *GeeksforGeeks*, 2.
- Hossain, A. G. (2019). *Player's Performance Prediction in ODI Cricket Using Machine Learning Algorithms*. Dhaka: IEEE.
- Indeed for employers. (2020). What Are Project Deliverables? *Indeed*.
- Jain, D. (2023). We Have a Clear Winner: Cricket Scoring Sheet Vs CricHeroes Online Cricket Scoring App. *Medium*, 23-23.
- Jayanth, S. B. (2018). A team recommendation system and outcome prediction for the game of cricket. *Journal of Sports Analytics*, 263-273.
- Jr., J. F. (2017). Entity-Relationship Diagrams. *Springer Link*, 1-10.
- Kachur, L. (. (2018). 5 Reasons Why Devs Love GitHub (And Microsoft Buys It). *Dzone* , 1-2.
- Kalpdrum Passi, N. P. (2018). Increased Prediction Accuracy in the Game of Cricket using Machine Learning. *arxiv*, 1-2.
- Kendall, P. (2023). Are there any size limits on the squad size for international cricket match series? *Satck Exchange*, 120-122.
- Kritika. (2022). Types of Software Testing. *GeeksForGeeks*, 2.
- Kumar, J. (2021). A study on Machine Learning Approaches for Player Performance and Match Results Prediction. *Cornell University*.
- Leeron Hoory, C. B. (2022). Agile Vs. Waterfall: Which Project Management Methodology Is Best For You? *Forbes Advisor*, 24-28.
- Lewis, S. (2023). prototyping model. *TechTarget*, 34.
- Liu, T. (2017). MockPlus. *What Are The Advantages of Prototyping? How to Make The Most of It?*
- Madur. (2021). 12 Backend Development Tools For Web Developers. *GeeksForGeeks*, 3.

- Maurya, R. (2020). Difference between Sequence Diagram and a System Sequence Diagram. *Digital Gyan*.
- Mehta, N. (2020). Let's Be Honest: Cricket is Dying a Slow, Yet Avoidable Death. *Medium*, 3.
- MELTZER, R. (2021). What Is Random Forest? *CareerFoundry*.
- Mohanta, V. p. (2020). Twenty20 Cricket – The Most Entertaining Cricket Format. *Indies Education*, 38-40.
- Monali Shetty, S. R. (2020). Monali Shetty; Sankalp Rane; Chaitanya Pandita; Suyash Salvi. *IEEE*.
- Naimishsahu08. (2021). Software Prototyping Model and Phases. *GeeksforGeeks*.
- Nayan Ranjan Das, S. G. (2023). Adoption of a ranking based indexing method for the cricket teams. *Science Direct*, 213.
- Nik. (2022). Introduction to Scikit-Learn (sklearn) in Python. *DataGy*.
- Okinyo, D. (2020). THE SILENT HERO Asiji is the face behind cricket development in Kenya. *The Star*, 5-7.
- Pal, S. K. (2021). Software Engineering | Phases of Prototyping Model | Set – 2. *GeeksforGeeks*.
- PortSwigger LTD. (2022, 05 19). *Burp Suite is the choice of security professionals worldwide*. Retrieved from portswigger.net: <https://portswigger.net>
- Rai, M. (2019). The Importance of Black Box Testing. *Happiest Minds Blogs*, 1.
- Rajesh, V. (2018). IPL 2018: Each franchise's key battles for a spot in the starting XI. *sportkeeda*, 13-19.
- Reshko, A. (2020). Top 7 Languages for Web App Development. *FortySeven*, 3-7.
- Shah, A. (2020). Intelligent Cricket Team Selection by Predicting Individual Players' Performance using Efficient Machine Learning Technique. *International Journal of Engineering and Advanced Technology*, 3406-3409.
- Sharma, R. (2020). Why Laravel is the Best PHP Framework to Use in 2020. *freeCodeCamp*, 1-9.
- Singh, V. (2022). Top 15 Web Development IDE. *hackr.io*, 2-9.
- T, N. (2020). Prototyping Model. *Binary Terms*, 4.
- Team Asana. (2022). A 6-step guide to requirements gathering for project success. *Asana* , 6.
- Thiraviam, B. (2023). How are ICC Rankings calculated? *CricIndeed*, 2-3.
- Tutorialspoint . (2022, March 11). *UML - Class Diagram*. Retrieved from www.tutorialspoint.com: <https://www.tutorialspoint.com/>
- Visual Paradigm. (2019). What is Activity Diagram? *Visual Paradigm*.
- Wigmore, T. (2014). Kenya cricket is dead and buried. *ESPN cricinfo*, 2.

- Williams, S. (2021). 10 Reasons Why Chrome is Better Than All Other Browsers. *TechArrival*, 3-6.
- Your Wise Information. (2022). Who decides the playing 11 in cricket? *YourWiseInformation*, 2-3.

Appendix

Appendix 1



Appendix 1: Gantt Chart

Appendix 2

Strathmore University
School of Computing and Engineering Sciences
Information Systems Project Proposal Assessment Guide

Table 2: Student Details

Student Number(s)	
Working Title:	

Table 3: Marking Guide

Evaluation Points	Weight	Score	Notes
Title page: Informative, concise and appropriate	2 pts		
Abstract To have background, problem, solution, methodology (approach data and tools) outcomes and expectations <i>Check on Completeness and correctness</i>	3 pts		
Introduction Background (2) <i>A clear illustration of issue, context and audience</i> Problem Statement (2) <i>Pain points, audience, who is affected and how solution comes in to fix the pain. What facts support this</i> Objectives (S.M.A.R.T and Linked to Problem Statement) (3) Research questions (3) <i>Alignment of questions with objectives</i> Justification (2) <i>Should be research supported.</i> Scope of Project (2)	(17 pts)		

<p><i>Specify boundaries of people process, HW/SW, data etc</i></p> <p>Limitations (1.5)</p> <p><i>Challenges Expected</i></p> <p>Delimitation (1.5)</p> <p><i>To do to counter anticipated challenges</i></p> <p><i>Check for correctness, completeness and citation of work</i></p>			
<p>Literature Review/Related Work</p> <p>Literature objectives mapping as aligned with research questions (2)</p> <p>Critique of content adequacy of What it is, how it presents, its implications, Citations of content align with work (4)</p> <p>Review of at least 3 systems comprehensively the working behind it (2)</p> <p>Gaps identification, analysis relative to the proposed solution (3)</p> <p>Conceptual Framework clear to communicate how it works, data flows, processing, actors (2)</p> <p>Describe input process output storage boundaries Emerging technologies contextualization (5)</p>	(18 pts)		
<p>Intended Approach/ Methodology</p> <p>Research Design (2) <i>experimental, casual etc to determine type of data to be used, Variables etc</i></p>	(9 pts)		

Research Methodology (5) <i>Methodology (1), Correct process (1), Design and Development tools (1), Research Paradigm (2)</i> Deliverables and milestones (2) <i>Examinable bits from ideation</i> <i>Proposal, design, test cases documentation doc</i> <i>Proof of concept- modules</i>			
Proposal Presentation Table of Contents and List of Figures (2) Are relevant references provided and formatted correctly? (1) Is there a clear and proper use of language? (1) Effective report structure (chapters and sections) and layout (2)	(6 pts)		
Total Marks	55		

Verdict (Please tick)	Accepted	Reject
-----------------------	----------	--------

Comments (Especially if verdict is reject)
