HENRY MBURU KURIA - IF/17/17

INFORMATICS PROJECT II

INF 490

# FOOTBALL PREDICTIONS WEBSITE

### Abstract

The gambling industry in Kenya has been on the rise recently. One indicator is the large number of bookmakers operating in the country using various website platforms such as Sportpesa, Betway and Mozzart bet.

This has led to a huge population of Kenyans becoming gamblers. A gambler is a risk taker who spends most of their time researching the most accurate bets to place a wager on. A good source of betting information could be the difference between a successful gambler and one who is not.

This study looks at football predictions, one of the most sought-after pieces of information by gamblers. Specifically, it seeks to explore the different online platforms that are used to disseminate this information and their down sides. This document further goes on to design an information system that improves on these downsides to provide a most efficient resource of said information

Machine learning is explored as a way that is used to come up with football predictions and eventually the information system that is designed in this study uses the results of machine learning and an optimal interface to present betting information to gamblers.

Table of Contents

[FOOTBALL PREDICTIONS WEBSITE I](#_Toc78893656)

[Abstract I](#_Toc78893657)

[List of Figures III](#_Toc78893658)

[List of Tables IV](#_Toc78893659)

[CHAPTER ONE: INTRODUCTION 1](#_Toc78893660)

[Introduction 1](#_Toc78893661)

[Background Information 2](#_Toc78893662)

[Brief History on Internet Gambling 2](#_Toc78893663)

[Football Betting in Internet Gambling 3](#_Toc78893664)

[Problem Statement 5](#_Toc78893665)

[Aim and Objectives 6](#_Toc78893666)

[Scope of the Study 7](#_Toc78893667)

[CHAPTER TWO: LITERATURE REVIEW 8](#_Toc78893668)

[Introduction 8](#_Toc78893669)

[Theoretical Framework 10](#_Toc78893670)

[Social Media Fatigue and Information Overload 10](#_Toc78893671)

[Accuracy of Betting Tips 11](#_Toc78893672)

[Quantity of Information 11](#_Toc78893673)

[Conceptual Framework 13](#_Toc78893674)

[VARIABLES ISOLATED FROM THE LITERATURE 13](#_Toc78893675)

[CHAPTER 3 - SYSTEMS METHODOLOGY 17](#_Toc78893676)

[Methodology and Approach of the Project 17](#_Toc78893677)

[Use of Machine Learning to make football predictions 20](#_Toc78893678)

[Systems Analysis 29](#_Toc78893679)

[Feasibility Study 29](#_Toc78893680)

[Requirements Gathering 30](#_Toc78893681)

[Requirements Review 31](#_Toc78893682)

[CHAPTER FOUR: SYSTEMS DESIGN AND IMPLEMENTATION 32](#_Toc78893683)

[Architectural Design 32](#_Toc78893684)

[Model View Controller Framework 32](#_Toc78893685)

[Cloud Deployment Architecture 33](#_Toc78893686)

[Process Flow Diagrams 34](#_Toc78893687)

[Interface Design 35](#_Toc78893688)

[Home Page 35](#_Toc78893689)

[Single Match 36](#_Toc78893690)

[Betslips 37](#_Toc78893691)

[Sort Betslips 38](#_Toc78893692)

[Authentication 39](#_Toc78893693)

[Search Results 41](#_Toc78893694)

[Database Design 42](#_Toc78893695)

[Entities 42](#_Toc78893696)

[Entity Relationship Diagrams 47](#_Toc78893697)

[Tools and Technologies Used 49](#_Toc78893698)

[Testing 50](#_Toc78893699)

[Deployment 51](#_Toc78893700)

[Conclusion and Recommendations 52](#_Toc78893701)

[Recommendations 52](#_Toc78893702)

[REFERENCES 53](#_Toc78893703)

# List of Figures

[Figure 1 winning teams and the accuracy of their ML models 12](#_Toc78893704)

[Figure 2 Variables isolated in the study 13](#_Toc78893705)

[Figure 3 gambler satisfaction against quantity of tips 14](#_Toc78893706)

[Figure 4 gambler satisfaction against accuracy of tips 15](#_Toc78893707)

[Figure 5 gambler satisfaction against ease of access of tips 16](#_Toc78893708)

[Figure 6 bar plot comparing viewers of telegram posts and the number of posts 18](#_Toc78893709)

[Figure 7 sample of the England match dataset 20](#_Toc78893710)

[Figure 8 confusion matrix for the England dataset ML model 26](#_Toc78893711)

[Figure 9 how to interpret a confusion matrix 26](#_Toc78893712)

[Figure 10 Evaluation metrics for the England dataset 'Both Teams To Score' ML model 27](#_Toc78893713)

[Figure 11 ROC curve for the England dataset 'Both Teams To Score' ML model 28](#_Toc78893714)

[Figure 12 django MVC framework 32](#_Toc78893715)

[Figure 13 Cloud Deployment Architecture 33](#_Toc78893716)

[Figure 14 Authentication Process 34](#_Toc78893717)

[Figure 15 Home Page 35](#_Toc78893718)

[Figure 16 Single Match Page 36](#_Toc78893719)

[Figure 17 Betslips Page 37](#_Toc78893720)

[Figure 18 Sort Betslips Page 38](#_Toc78893721)

[Figure 19 login page 39](#_Toc78893722)

[Figure 20 sign up page 40](#_Toc78893723)

[Figure 21 Search Results Page 41](#_Toc78893724)

[Figure 22 User Entity 42](#_Toc78893725)

[Figure 23 Match Entity 42](#_Toc78893726)

[Figure 24 Betslip Entity 43](#_Toc78893727)

[Figure 25 Competition Entity 43](#_Toc78893728)

[Figure 26 Country Entity 44](#_Toc78893729)

[Figure 27 Match Result Tip Entity 44](#_Toc78893730)

[Figure 28 Both Teams To Score Tip Entity 45](#_Toc78893731)

[Figure 29 Over/Under Tip Entity 45](#_Toc78893732)

[Figure 30 Tip Abstract Class 46](#_Toc78893733)

[Figure 31 Entity Relationship Diagram 47](#_Toc78893734)

[Figure 32 Database Entity Relationship Diagram 48](#_Toc78893735)

# List of Tables

[Table 1 telegram post viewers considering accuracy of tips 22](#_Toc78817103)

[Table 2 how easy it was to find tips in a telegram channel 24](#_Toc78817104)

[Table 3 Testing table 58](#_Toc78817105)

# CHAPTER ONE: INTRODUCTION

## Introduction

In the gambling industry, having a good source of predictions can be the difference between winning bets and losing bets. In football, there are many matches on a daily basis which presents more opportunities for gamblers to earn. Some people gamble as a hobby however it is not uncommon to find a gambler who gambles for a living.

What is gambling? Oxford dictionary defines gambling as playing games of chance or money. A second definition from the dictionary goes on to define gambling as taking risky action in the hope of a desired result. [Britannica Encyclopedia](https://www.britannica.com/topic/gambling) “Gambling, the betting or staking of something of value, with consciousness of risk and hope of gain, on the outcome of a game, a contest, or an uncertain event whose result may be determined by chance or accident or have an unexpected result by reason of the bettor’s miscalculation.”

Making successful bets is not easy. A gambler has to make a prediction on the outcome of a match based on many factors such as the results of the previous meetings between the two teams, the players in a team, the position of the team in the league table, the attacking strength of each of the two teams and the defending strength of the teams among many other factors.

For the successful gambler, it takes time to analyze a football match as many factors have to be taken into consideration as seen from the examples above. This means that gamblers are only able to analyze a few matches at a time. This has led to the need of an information system that provides both knowledge and variety in that knowledge, to aid a gambler in the decision-making process.

## 

## 

## 

## Background Information

### Brief History on Internet Gambling

Internet gambling is gambling on, or by means of the internet. ‘Internet gambling’, ‘online gambling’, ‘e-gaming’, ‘remote gambling’ and ‘interactive gambling’ are interchangeable terms used to describe gambling that occurs via the internet, interactive television or mobile phone (Williams et al., 2012). In 1995, a few sites started offering casino gambling games online without real money being wagered. The first case of money actually being wagered over the internet by the general public appears to be the online purchase of lottery tickets from the International Lottery in Liechtenstein for a manual drawing that occurred on October 7, 1995(Romney, 1995).

In 1997 there were fifty to sixty Internet casinos in operation, most based in the Caribbean, which earned approximately $300 million to $350 million. By 2000 an estimated six hundred to seven hundred sites were operating and revenues approached $2 billion ("Internet Gambling." Gambling: What's at Stake? Retrieved April 15, 2021 from Encyclopedia.com: <https://www.encyclopedia.com/reference/social-sciences-magazines/internet-gambling>). By 2007 the American Gaming Association (AGA) estimated in the fact sheet “Internet Gambling” (2007, http://www.americangaming.org/Industry/factsheets/issues\_detail.cfv?id=17) that about two thousand online casinos were estimated to exist.

### Football Betting in Internet Gambling

Football betting is just one of the many forms of gambling. Here, we look at some of the popular types of bets in football gambling. By understanding these bets, we are able to understand which kind of information gamblers look for in a source of betting information.

The following types of bets are offered by betway.co.ke(<https://www.betway.co.ke/betting-rules-and-tips>), a gambling website.

* **Full Time Result**

The most popular and straightforward bet you can find. You’re simply asked to bet on the match winner (1 for the Home Team, 2 for the Away Team) or a draw (X).

* **Correct / Correct Score**

For Correct Score betting, predict the actual score line after 90 minutes of play, i.e. full time. For example: Liverpool 2 – 2 Real Madrid

#### **Half-Time/Full-Time**

For Half Time/Full Time betting, predict the outcome at half-time and at the end of regulation 90 minutes play (e.g., half-time Draw, full-time Home Team).

#### **Half Time Result**

Select the match status at half time (e.g., Home Team winning the first half).

#### **Over/Under**

#### You bet on the total number of goals that will be scored by both teams, in a match. The most common limit is 2,5 goals (over 2.5 / under 2.5). So, to win an Over 2,5 bet you need 3+ goals and 2 or less for an Under 2,5 bet.

#### **2 Way**

#### This is like a full-time result bet but there is no option for the draw. You can bet on either team to win and if it’s a draw you will get your stake back.

#### **Double Chance**

#### The Double Chance bet allows you to cover two of the three possible win/draw/win (1X2) outcomes in one bet.

#### Double Chance bets have 3 selections:

* 1/X - The home team will win or the match will end in a draw
* X/2 - The away team will win or the match will end in a draw
* 1/2 - Either the home or the away team will win the match

#### **Goal/No Goal**

#### For the Goal/No Goal Market, “Goal” is for both teams to score, “No Goal” is for either or both teams not to score.

How much the gambler wins is dependent on two factors: How much the gambler waged and the odds for an event taking place. The odds of an event are calculated by the bookmaker. They represent the probability of an event taking place. If an event takes place, the gambler earns his wager multiplied by the odds.

If the event does not take place, the gambler loses the bet and the wager placed on it.

A gambler can place more than one bet using the same wager. The collection of bets placed under one wager is referred to as a bet slip/accumulator (<https://www.betway.co.ke/betting-rules-and-tips>)

## Problem Statement

A gambler's satisfaction in a source of betting information is impacted by the accuracy of the information on the source, the ease of accessing information on the source and the quantity of information on the source.

Who are tipsters? Tipsters are sports betting experts that offer recommendations concerning betting opportunities, usually in exchange for money (Gonzalez, 2020). (Miller et al., 2016) defines tipsters on social media as “accounts that offered advice and tips about bets and gambling that are frequently paid affiliates of gambling sites”

In recent years, tipsters have taken to social media as a means of providing tips to gamblers and also as a way to increase their popularity as reputable sources of accurate predictions. This is seen in tipster communities which are made up of people who mainly follow tipsters (Miller et al., 2016). This has led to a large outburst of many sources of betting tips. Due to the many tipsters on social media, there is a lot of inaccurate betting information out there.

Another unintended consequence of the large number of tipsters online is disorganization of betting information. Gamblers try to find information regarding many matches and many predictions. A match can have ten or more predictions and each day can have more than twenty matches. Due to the vastness of information that gamblers search for, it is easy for a source of betting tips to have disorganized information which makes it difficult for gamblers to access the information that they need.

## Aim and Objectives

The aim of this study is to investigate accuracy, quantity and ease of access of information as the problems that impact how satisfied a gambler is with a source of betting information and to develop an information system that addresses these problems

Objectives:

1. To explore the different platforms that are used to present betting information to gamblers and their downsides.
2. To explore the different ways that are used to make football predictions
3. To develop an information system that provides betting information to gamblers and that which has an interface that addresses the problems on other sources of betting information

## Scope of the Study

This study will focus on the problems that gamblers face when using Twitter and Facebook social media platforms as sources of betting information. It will also look at machine learning as a means of making football predictions.

This study will also cover the detailed description and design of an information system that will solve the problems that gamblers face.

The population under study is the subscribers of various betting communities on telegram, Facebook and twitter which provide betting information and the tipsters that maintain these communities.

# CHAPTER TWO: LITERATURE REVIEW

## Introduction

Gamblers are always looking for betting tips and tipsters are always looking for channels to distribute tips.

Social media is one of the tools that are used by tipsters to distribute betting information to gamblers. In this document, focus is put on Facebook and twitter as social media platforms used in distributing gambling information

“Social media is synonymous with community. People want to connect; they want to talk. They want to share. They belong to whatever community they want-no matter how big or small” (Lincoln, 2009). Social media has allowed new communities to form, related to gambling, sometimes around specific apps and software, some around particular affiliates, tipsters and content-producers, and others around problems and issues, including problematic gambling (Miller et al., 2016).

#### Twitter

One form of social media used by tipsters to distribute betting information is microblogging, a concept commonly associated with Twitter, the most widely used microblogging site in the world [Burton, Dadich & Soboleva,2013].

Twitter allows users to post short text messages to individuals who have chosen to ‘follow’ the sender and followers are able to actively engage and forward or ‘retweet’ other’s messages [Burton, Dadich & Soboleva, 2013]. Launched in 2006, Twitter has 302 million daily users, who collectively send over 500 million tweets daily [Twitter, 2015]. Of these active users, 80% use their mobile phone as a medium to tweet and/or retweet [Twitter, 2015].

The largest form of promotion via Twitter was tweets which prompted individuals to place bets (Thomas et al, 2015, pg 35). This is the main type of content that tipsters distribute, information that encourages a gambler to place a bet

(Miller et al., 2016) identified 877 twitter accounts as dedicated to producing content promoting gambling. They sent over 78,000 tweets during the period of the study, two per minute. The largest community, of 140,000 members, tended to follow tipsters and affiliates rather than main bookmakers.

#### Facebook

Facebook is the most popular social networking site in the world (Freeman et al. 2014). As of March 2015, 1.44 billion users accessed the site at least monthly and 936 million accessed the site daily (Facebook, 2015).

From November 2007, Facebook embraced companies and commercial brands developing their own Facebook pages, and thus their own online identities [Freeman et al. 2014]. Companies

can ‘post’ images, videos, links, offers, competitions and a range of other digital media to

their page timelines [Freeman et al. 2014]. When consumers ‘like’ a company page to receive timeline updates, any consumer engagement with company pages may appear in the news feed of another’s Facebook [Freeman et al. 2014]. Thus, companies are able to effortlessly spread their marketing messages and brand images across multiple social networks [Griffiths & Casswell, 2010;

Tipsters, like companies and online brands, use Facebook to share information with their followers. They are able to share betting tips to gamblers. Facebook also enables tipsters to manage their audience better and get feedback on their performance by looking at their growth in followers. The most used gambling promotions on Facebook were prompts to bet and tipping (Thomas et al, 2015, pg 39).

A case study on three football betting tips Facebook pages with more than 79,500, 79,000 and 27,000 likes each showed that they each published 5 -20 posts per day, attracting 10-200 likes and 10-250 comments per post. These three pages were similar in terms of content and purpose (Miller et al., 2016).

## Theoretical Framework

How satisfied a gambler is with a source of betting information can be referred to as customer satisfaction. Customer satisfaction is “the perception of the customer as a result of consciously or unconsciously comparing their experiences with their expectations” Thomassen (2003, p. 69). The definition of Zeithaml and Bitner (2003, p. 86) is slightly different from that of Thomassen: “Satisfaction is the consumer fulfillment response. It is a judgement that a product or service feature, or the product or service itself, provides a pleasurable level of consumption-related fulfillment”.

For this study, Zeithaml and Bitner definition is to be used since a gambler's satisfaction on a source of tips is theoretically based on the accuracy of tips, ease of access of the tips and the quantity of information on the source. Below we explore the different factors that lead to gambler **dissatisfaction** in various betting information sources online.

### Social Media Fatigue and Information Overload

This factor directly impacts the ease of access of information on a source of betting information

Social media fatigue is defined as a user’s tendency to back away from social media participation when s/he becomes overwhelmed with information. Privacy concerns and confidence have the greatest predictive value for social media fatigue (Bright et al., 2015). Social media fatigue leads to individuals’ discontinuous usage of social media (Fu et al, 2020).

Modern social media are becoming overloaded with information because of the rapidly-expanding number of information feeds (Feng et al, 2015). Survey results on a UK-based online questionnaire investigating aspects of usage of user-generated media (UGM), such as Facebook, LinkedIn and Twitter which attracted 587 participants showed that two thirds of Twitter-users have felt that they receive too many posts, and over half of Twitter-users have felt the need for a tool to filter out the irrelevant posts (Bontcheva et al., 2013).

When tipsters use social media to distribute betting tips, they are in constant competition with each other for the attention of users. The chances of an individual to share a message is proportional to the fraction of its neighbors who shared it with him/her, which is a result of competition for attention (Feng et al, 2015). This leads to information overload on the users.

This means that despite many tipsters using social media to distribute tips, the gamblers may suffer from social media fatigue and information overload. This would lead them to cease from using social media and thus disregard social media as an adequate source of betting information.

A website on the other hand has information from only one source, the website owner who in this study is a tipster. When gamblers visit one tipster website, they view tips from only that tipster and the chances of information overload are reduced. On a website, the gamblers will also not suffer from social media fatigue.

### Accuracy of Betting Tips

(BBC, 2016) estimates that around half of online tipsters are affiliated with a betting company. This means that when tipsters post their tips on social media, they include a link to the betting company's site. What is rarely made clear is that the tipster has a financial arrangement with that betting company, one which means the tipster either receives a fixed sum for every person who creates an account with the firm, or more popularly, takes a cut of around 30% of the punter's losses. If you lose sh1000 on a bet, the tipster gets to pocket sh300 for themselves, even if it was their bad advice that lost you the money (BBC, 2016). This shows that tipsters can intentionally deceive gamblers.

Another issue leading to inaccuracy of betting tips is the presence of many scammers online who claim to be tipsters. (Jonathan, 2020) explores a strategy to use as a tipster to retain gamblers without having any knowledge to make football predictions. The tipster flips a coin to make a prediction. The tipster then gives half the gamblers one side of an event and the other half the other side of the same event. (Jonathan, 2020) goes ahead to say that most tipsters use this model of marketing to reach and keep more gamblers loyal to their tipster service. This kind of marketing strategy leads gamblers to making wrong decisions

### Quantity of Information

Given the many matches each day and many markets for predictions per match, it is impossible for a human to make predictions for all matches in a day. A gambler could want tips on a specific match to be played that day, but the tipster may lack the information.

To counter this problem, machine learning could be used. Machine learning is a branch of artificial intelligence that systematically applies algorithms to synthesize the underlying relationships among data and information (Khanna, 2015). In 1959, Arthur Samuel described ML as the “field of study that gives computers the ability to learn without being explicitly programmed” (Samuel 1959)

Artificial intelligence (AI) research has become prominent in both academia and industry. With this, an interest in AI's ability to make sound decisions when compared to human decision making has grown. Predicting the outcome of sporting events has traditionally been seen as a difficult task, due to the complex relationships between variables of interest (Pretorius & Parry, 2016).

In 2017, a Soccer Prediction Challenge that revolved around predicting the outcomes of future soccer matches was held for the general public. The fundamental research question of the 2017 Soccer Prediction Challenge was the following: “To what extent is it possible to predict the outcome of a soccer match, given commonly available match data?” The competition’s task was to use machine learning to predict the outcome of future soccer matches. Using only the provided data, the Challenge participants had to develop machine learning models in order to predict the outcome of 206 future matches that took place after the submission deadline. Thus, when the participants submitted their predictions, the outcomes for these matches were not known to anyone. The Database, the 2017 Soccer Prediction Challenge and its results are described in Dubitzky et al.’s article entitled “The Open International Soccer Database for Machine Learning” (Dubitzky et al. [2018](https://link.springer.com/article/10.1007/s10994-018-5763-8#ref-CR17)).

A sample of the results is as follows:

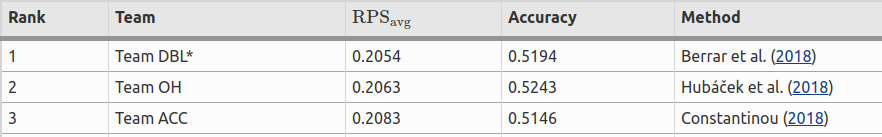


Figure 1 winning teams and the accuracy of their ML models

The emphasis on this challenge for this study is to show that indeed, machine learning can be used to make football predictions. Further publications show the increase in accuracy based on factors beyond the scope of this study.

In (Pretorius & Parry, 2016), A classification algorithm was employed to predict match outcomes in the 2015 Rugby World Cup. The performance of this model was compared to aggregate results from Super-Bru and OddsPortal. The machine learning based system achieved an accuracy of 89.58%. These results indicate that for rugby, over the limited period of a specific tournament, the evidence was **not** strong enough to suggest that a human agent is superior in terms of accuracy when predicting match outcomes compared to a machine learning approach.

The focus of exploring machine learning as a way of making football predictions in this study is to counter the problem of large scale decision making. By use of machine learning, tipsters could provide tips on many matches with high automation and accuracy. This would theoretically solve the problem of quantity in information sought after by gamblers

## Conceptual Framework

### VARIABLES ISOLATED FROM THE LITERATURE

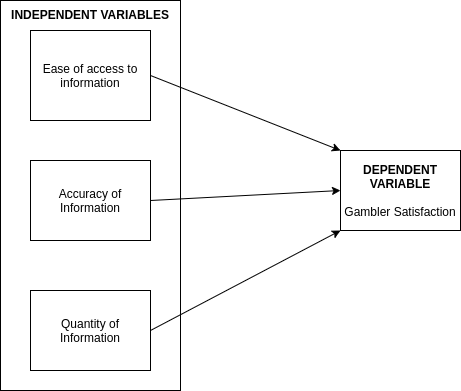


Figure 2 Variables isolated in the study

#### Gambler Satisfaction in relation to Quantity of Tips

The diagram below shows the expected effect on gambler satisfaction that increasing the quantity of tips on a source of betting information has. Gambler satisfaction would be expected to increase with increase in quantity of tips until a point reaches where it plateaus. It is expected to plateau since a gambler usually wants tips on specific matches and when they find these tips, they are no longer interested in other matches, thus, even though quantity of tips increases for other matches, their satisfaction levels remain the same.

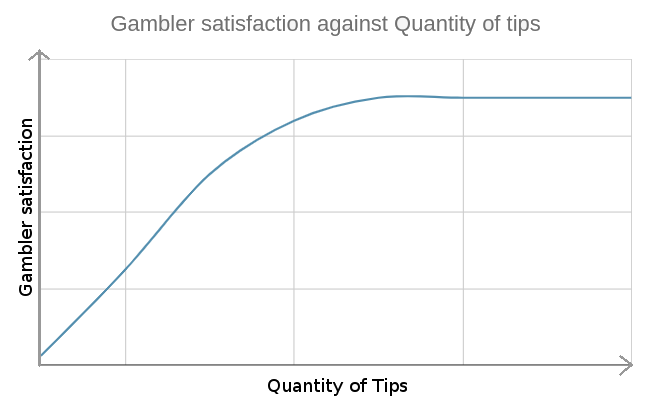


Figure 3 gambler satisfaction against quantity of tips

#### Gambler Satisfaction in relation to Accuracy of Tips

For accuracy, gambler satisfaction would not be expected to decrease since they will always win more when the source is accurate. Below shows the expected relationship between gambler satisfaction and accuracy of betting tips

As the accuracy of the tips increases, gambler satisfaction is expected to increase. Gambler satisfaction is directly proportional to accuracy of tips provided

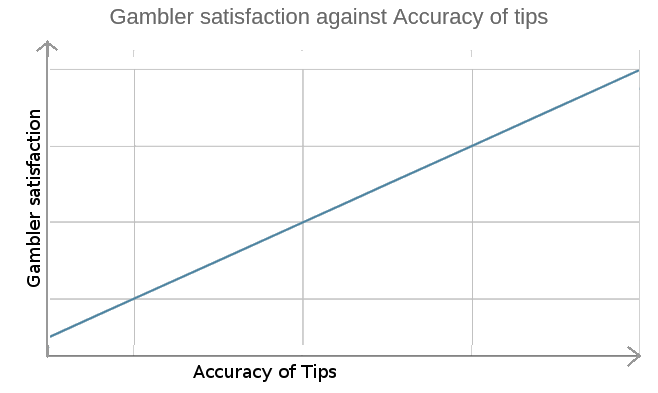


Figure 4 gambler satisfaction against accuracy of tips

#### Gambler Satisfaction in relation to Ease of access to tips

The easier a source of betting information is to navigate and find information on, the more satisfied a gambler is with that source. This satisfaction is, however, expected to plateau since gamblers eventually learn how to get the information they need from the source and the focus now becomes about the accuracy and quantity of information on that source.

The diagram below shows a depiction of this relationship

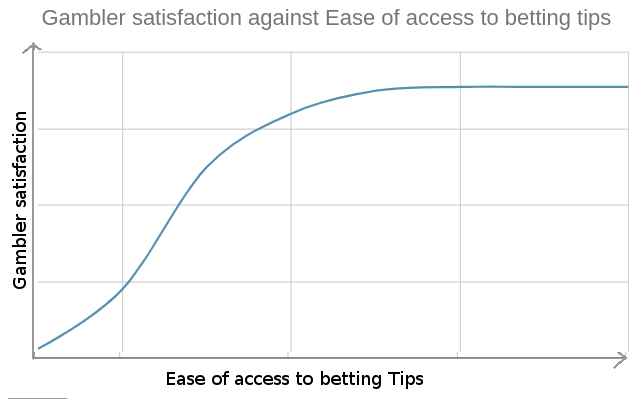
****

Figure 5 gambler satisfaction against ease of access of tips

# CHAPTER 3 - SYSTEMS METHODOLOGY

## Methodology and Approach of the Project

To conduct this study, the behavior of gamblers in a telegram channel which provided betting tips to gamblers and the posts made to the channel were observed. The channel had 200 subscribers and had been operational for six months. The measure for satisfaction used here was the number of views that a post in that channel got.

**Questions asked**

1. Did the telegram channel get more viewers when they were more accurate?

To measure this, the accuracy of the channel every day was recorded for a period of three months along with the daily viewership. The data was then analyzed to find the relationship between the accuracy and the viewership of the channel

Below shows a sample of data collected to answer this question

Table 1 telegram post viewers considering accuracy of tips

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Date  (yy-mm-dd) | total number of tips posted | total number of correct tips | Accuracy in percentage  (correct tips / number of tips posted) | Total viewers |
| 2020-08-04 | 25 | 21 | 84% | 197 |
| 2020-08-05 | 42 | 35 | 83% | 200 |
| 2020-08-06 | 38 | 30 | 78% | 245 |
| 2020-08-07 | 47 | 36 | 76% | 158 |
| 2020-08-08 | 38 | 34 | 89% | 198 |

1. Did the quantity of tips posted have an impact on the viewership?

By using the data from table 1, the relationship between total number of tips posted and total viewership could be established by plotting a graph.

For the sample above the following was observed

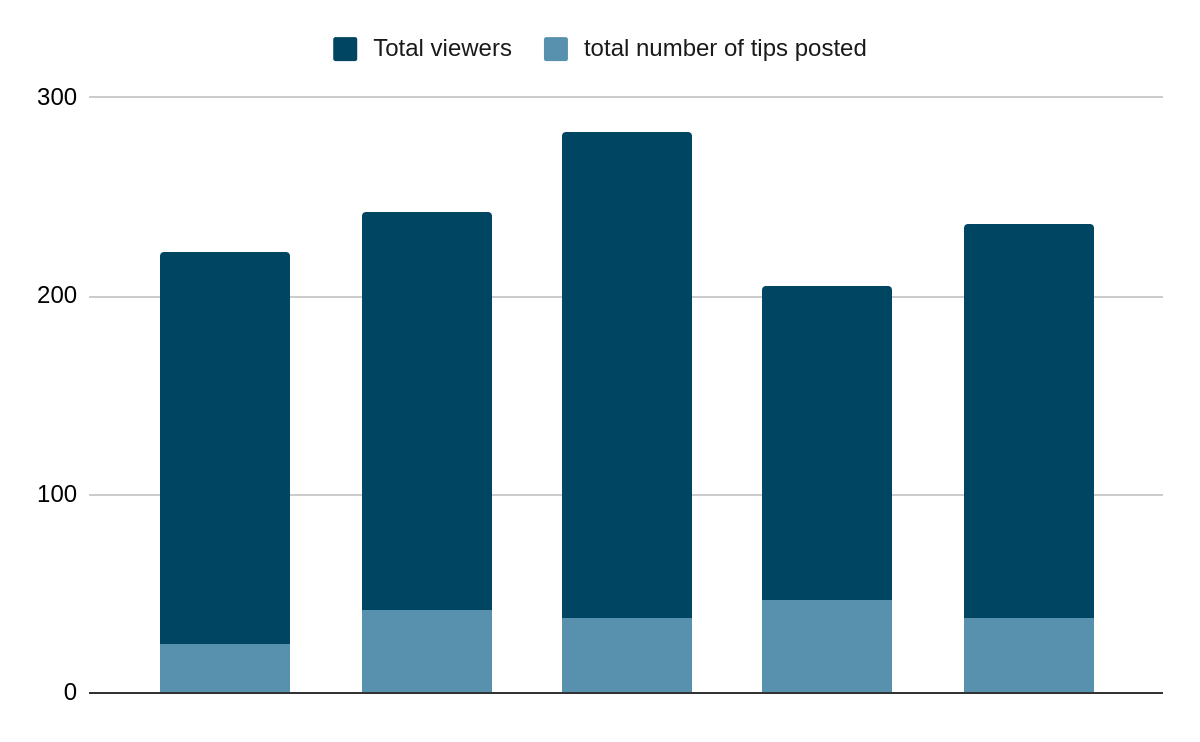


Figure 6 bar plot comparing viewers of telegram posts and the number of posts

1. How easy was it to find a specific tip in the telegram channel?

To answer this question, questionnaires were given to the 200 subscribers of the telegram channel asking them how likely they were to return the next day based on their experience in finding tips on a specific date.

Below shows a sample of the questionnaire used to collect data necessary to answer this question.

Table 2 how easy it was to find tips in a telegram channel

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
| Date | number of tips being looked for | number of tips found | Time taken to find the tips  (minutes) | How easy was it to the tips?  (1 - very difficult, 10 - very easy) | How likely are you to come back tomorrow to find other tips?  (1 - very unlikely, 5 -very likely) |
| 2020-08-04 | 5 | 5 | 2 | 4 | 2 |
| 2020-08-05 | 10 | 8 | 2.5 | 6 | 4 |
| 2020-08-06 | 6 | 5 | 3 | 3 | 2 |
| 2020-08-07 | 5 | 5 | 2 | 5 | 2 |
| 2020-08-08 | 10 | 8 | 3.5 | 2 | 3 |

Specifically comparing the results of the questionnaire and the number of posts made in a day in table 2, it could be established if the subscribers suffered from social media fatigue. We could also get to measure how satisfied a gambler was with the channel

### Use of Machine Learning to make football predictions

To make the football predictions, machine learning is applied. In this section, focus is put on how predictions were made for matches in England, where the competitions included; English Premier League, Championship, League One, League Two and the FA Cup.

#### Data Collection

The primary source of the data was the internet. The data collected was that of previous matches and upcoming matches (fixtures). The previous matches would be necessary to create the machine learning models while the fixtures would be the matches that the predictions would be made on.

The time variance of the data ranged from the year 1995 to the year 2021. The England dataset had a total of 12,150 previous matches and 675 fixtures.

#### Data Exploration

After collection of data, the dataset collected had the following columns

Away Team - The away team in the match

Home Team - The home team in the match

Full Time Home Goals (FTHG) - Goals that the home team scored

Full Time Away Goals (FTAG) - Goals that the away team scored

Date - The date that the match was played

League - The competition in which the match was being played for example, the English Premier League or the FA Cup

Match info - Whether the match ended at full time or extra time

season - The season that the match was played in for example; 2020-2021

starting time - The time that a match started

Below shows a sample of the data in the England dataset, showing the various columns



Figure 7 sample of the England match dataset

#### Feature Engineering

Features answer questions about a dataset. According to <https://www.datarobot.com/wiki/feature/>, A [feature](https://www.datarobot.com/wiki/feature/) is a measurable property of the object you’re trying to analyze. In datasets, features appear as columns. The quality of the features in your dataset has a major impact on the quality of the [insights](https://www.datarobot.com/wiki/insights/) you will gain when you use that dataset for [machine learning](https://www.datarobot.com/wiki/machine-learning/). Martin Heller describes a feature as an individual measurable property or characteristic of a phenomenon being observed (<https://www.infoworld.com/article/3394399/machine-learning-algorithms-explained.html#:~:text=Recall%20that%20machine%20learning%20is,classification%2C%20regression%2C%20etc>)

For each match in the England dataset, the following features were engineered.

* **Team Attacking Strength**

This is the attacking strength of a team. It is the ratio between the league’s and the team’s averages (“How to predict the score in football betting”, Bettingwell )

***Average number of goals scored at home throughout all seasons in the dataset / Average number of goals scored away throughout all seasons in the dataset***

* **Team Conceded Goals**

This is the number of goals scored against a team over all matches in the dataset.

For example, in a match between Manchester United and Arsenal which ended at 2 - 1, Manchester United would have conceded 1 goal, the goal that was scored by Arsenal.

* **Team Draws**

These are the draws that a team has had in all the matches played throughout all seasons in the dataset

* **Team Defending Strength**

This is the defending strength of a team. It is the ratio between the league’s and the team’s averages.

***Average number of goals conceded at home throughout all seasons in the dataset / Average number of goals conceded away from home throughout all seasons in the dataset***

**Source:** (“How to predict the score in football betting”, Bettingwell )

* **Team Goal Difference**

Oxford dictionary ([https://www.oxfordlearnersdictionaries.com/definition/english/goal-difference](https://www.oxfordlearnersdictionaries.com/definition/english/goal-difference#:~:text=%2F%CB%88%C9%A1%C9%99%CA%8Al%20d%C9%AAfr%C9%99ns%2F-,%2F%CB%88%C9%A1%C9%99%CA%8Al%20d%C9%AAfr%C9%99ns%2F,position%20in%20the%20league%20table)) defines it as the difference between the number of goals that a team has scored and the number of goals that have been scored against them over a series of games, sometimes used to decide the team’s position in the league table.

***Team goal difference = team goals - team conceded goals***

* **Team Goals Scored**

This is the total number of goals scored by a team throughout all seasons in the dataset

* **Team Losses**

This the total number of losses that a team encountered throughout all seasons in the dataset

* **Team Match Average Goals**

This is the average number of goals that a team scored in all matches where they played against a specific team.

For example, consider the following theoretical matches which are used to calculate the Arsenal’s match average goals

Manchester United vs **Arsenal**: **2** - 0

Manchester United vs **Arsenal**: 1 - **1**

**Arsenal** vs Manchester United: **1** - 3

**Arsenal** vs Manchester United: **3** - 2

Manchester United vs **Arsenal**: 1 - **2**

In the above matches, the average match goals for Arsenal are:

(2 + 1 + 1 + 3 + 2) / 5

Which equals to 1.8 goals

* **Team Match Attacking Strength.**

This is the attacking strength of a team, but it is limited to matches where a team played against a specific team as in the example in calculating **Team Match Average Goals.**

The formula for calculating the same as that of calculating **Team Attacking Strength**. The data input is the only difference

* **Team Match Defending Strength**

This is the defending strength of a team, but it is limited to matches where a team played against a specific team as in the example in calculating **Team Match Average Goals.**

The formula for calculating the same as that of calculating **Team Defending Strength**. The data input is the only difference

* **Team Played Matches**

This is the total number of matches that a team played throughout all seasons in the dataset

* **Team Previous Matches**

For each team, its previous 6 match results were extracted and simplified to either a win, a loss or a draw.

* **Team Wins**

This is the total number of games won by a team throughout all seasons in the dataset

* **Potential Away Goals**

This is the number of goals that the away team is expected to score in any match. It is derived by:

***Away Team Attacking Strength X Home Team Defending Strength X Average goals scored in the entire dataset***

**Source:** (“How to predict the score in football betting”, Bettingwell)

**Potential Home Goals**

This is the number of goals that the home team is expected to score in any match. It is derived by:

*Home Team Attacking Strength X Away Team Defending Strength X Average goals scored in the entire dataset*

**Source:** (“How to predict the score in football betting”, Bettingwell)

* **Goal Goal**

The value of this feature is either true or false. This feature answers the question: did both teams score in a match? If both teams scored, i.e., 2 - 1, the value would be ‘true’ and if both teams did not score, i.e., 2 - 0, the value would be ‘false’

* **Full Time Result**

The value of this feature is in the form of ‘home win’, ‘draw’ or ‘away win’. For example, in a match where the final score was 2 - 1, the value of this feature would be ‘home win’, in a match where the final score was 1 - 1, the value would be ‘draw’ and lastly, in a match where the final score was 1 - 2, the value would be ‘away win’

* **Over/Under 1.5 goals**

The value of this feature is either ‘true’ or ‘false’

This feature answers the question: were the total goals in the match greater than or equal to 2 goals? For example; in a match where the final score was 1 - 0, this value would be false while in a match where the final score was 1 - 1, the value of this feature would be ‘true’.

* **Over/Under 2.5 goals**

The value of this feature is also either ‘true’ or ‘false’

This feature answers the question: were the total goals in the match greater than or equal to 3 goals? For example; in a match where the final score was 1 - 0, this value would be false while in a match where the final score was 2 - 1, the value of this feature would be ‘true’.

* **Previous Matches**

This feature tells the number of times that two teams have played against each other. For example, in a match between Manchester United and Arsenal, the two teams could have met before as follows:

Manchester United vs Arsenal

Arsenal vs Manchester United

Arsenal vs Manchester United

Manchester United vs Arsenal

This would be equal to 4 previous matches

#### Modelling

Due to the large scope of the types of predictions made, focus here will be on modelling for predicting whether both teams will score in a match. The dataset is also limited to that of England

The problem of predicting whether both teams will score in a match was viewed as a classification problem whereby according to the features of each match, a match can be classified as either in the class of Yes (both teams will score) or No (both teams will not score). Many classification algorithms exist but for this study, logistic regression was used.

*Why Logistic Regression?*

As with many classification problems, the results of the machine learning model determine the effectiveness of the algorithm used. The nature of machine learning is iterative and the more the iterations and feature engineering, the better or worse your results could be. As such, Logistic regression was chosen simply because it fit the problem faced.

Other classification algorithms such as decision trees and vector machines could also be used to achieve similar or even better results, but that was not tested as it was not among the objectives of this study.

##### Brief summary of Logistic Regression

It is a statistical and machine learning technique for classifying records of a dataset based on the values of the input fields. It tries to predict a categorical or discrete target field instead of a numerical field. In predicting whether both teams will score, the value of the target field is discrete i.e ‘Yes’ or ‘No’

Logistic regression not only predicts the class, but also the probability of a case belonging to a specific class.

In logistic regression ^***y* = P(y=x)**

Where ^*y* is the prediction, P is the probability, y is the case and x is the target field.

For example;

The probability of match\_x to be a case in the class ‘Yes’ for both teams to score prediction is:

Yes = P(match\_x=‘Yes’)

##### Evaluation Metrics Used

These are used to determine how good or bad a model is at making predictions. The following shows the different results of the different evaluation metrics used

*Confusion Matrix*

It shows the model's capability to correctly separate classes.

Below shows the confusion matrix for the model used to predict the England dataset.

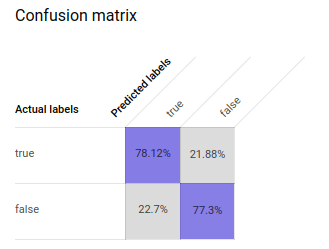


Figure 8 confusion matrix for the England dataset ML model

The percentage of classes predicted as ‘True’ in the entire dataset that were actually True were 78.12% while the percentage of classes predicted as ‘False’ in the entire dataset that were actually False were 77.3%. The percentage of classes predicted as ‘True’ in the entire dataset that were actually False were 22.7% while the percentage of classes predicted as ‘False’ in the entire dataset that were actually True were 21.88%.

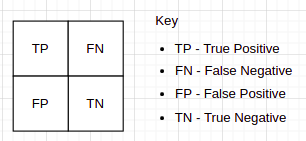
****

Figure 9 how to interpret a confusion matrix

From the confusion matrix, other metrics are also derived

**Precision**

This is the measure of the accuracy provided that a class label has been predicted.

Precision = True Positive / (True Positive + False Positive)

**Recall**

This is the true positive rate.

Recall = True Positive / (True Positive + False Negative)

The above two metrics are impacted by the threshold. (“Introduction to BigQuery”, Towards Data Science, retrieved 2021) explains threshold as the value from which the predictions of our model (which are values ​​between 0 and 1 that can be interpreted as probabilities that this observation is of class 1) will be for class 0 or for the class 1.

**F1 Score**

This is the harmonic average of the precision and the recall. It shows that a classifier has good recall and precision value.

F1 score = 2 x (precision x recall) / (precision + recall)

Below shows the results of the evaluation of the model used to predict the England dataset. A threshold of 0.5307 has been used.

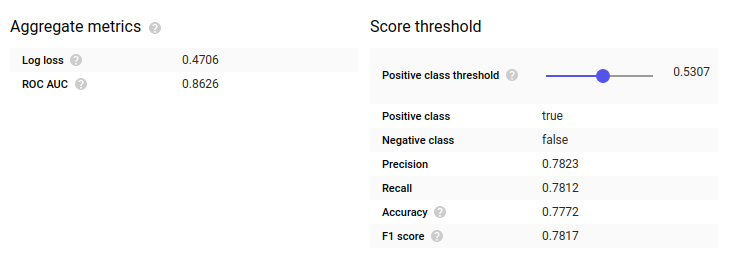


Figure 10 Evaluation metrics for the England dataset 'Both Teams To Score' ML model

Below shows a Receiver Operating Characteristic (ROC) curve for the model used to predict whether both teams will score in the England dataset. An ROC curve is a graphical visualization used to evaluate the predictive ability of a binary classification model.

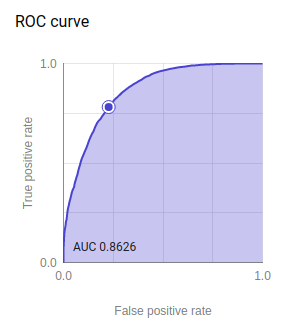


Figure 11 ROC curve for the England dataset 'Both Teams To Score' ML model

## Systems Analysis

### Feasibility Study

#### Legal Feasibility

The football predictions website developed at the end of this study will provide terms and conditions that users are supposed to abide by when using the website. These terms and conditions will serve to prevent any unintended legal challenges related to gambling.

The website will include a privacy policy that will explain the different scenarios that would require collection of user information and what the information would be used for. It is a requirement by law to provide this information to users before they agree to the collection of their data.

Due to the originality of the betting information on the website, it can engage the user to make a purchase of information without breaking any copyright law.

#### Technical Feasibility

The website is technically feasible up to a specific limit based on the funds available to run the project. The infrastructure needed to host and run the website is available on cloud resources and the only limiting factor is the funds.

Use of cloud infrastructure as a service enables the quick scaling, both up and down, of computing resources whenever needed for example in the case of a traffic spike to the website.

#### Economic Feasibility

With the use of cloud technology, the requirement to maintain the infrastructure upon which the website is deployed is outsourced to the cloud provider. This reduces the costs of maintenance of the website substantially.

The website is also economically feasible since it promises an eventual return in investment. This is in the form of the monetary value of the sale of part of the information on the website. This will generate revenue that will be diverted to cater for costs arising from its maintenance

Due to the high automation in the website, it will require very little labor to run. This ensures that there are minimal costs in securing human resource.

Another aspect of economic feasibility is that the website is to display ads to visitors. The ad providers considered are Google Ads and AdPushup. This is a passive form of earning revenue via the website since no input is required, just website traffic.

#### Operational Feasibility

The website has the flexibility to add more automation. This allows for automation of future processes. This allows for very little human intervention and thus cuts on the manpower needed to operate it. It also reduces the complexity involved in running the website

Information on the website is added in batches. This also reduces the amount of work needed to run the website

### Requirements Gathering

The method that was used to collect requirements was focus groups. A group of ten gamblers was invited to five focus group sessions to brainstorm the features that would be needed in the proposed information system. The population was made up of 6 males and 4 females all of whom were between the ages 18 and 26 years old. The members of this population were all familiar with gambling and 8 of them were active gamblers. They were chosen because they had used other information systems that provided tips to gamblers and they understood the psychology of gamblers.

Five sessions were conducted and, on each session, a prototype of the proposed information system was used. The members of the group would then give suggestions and improvements on the prototype until after the fifth meeting when the information system was deemed to have addressed the three variables of this study and was considered ‘acceptable’ to the public.

### Requirements Review

Out of these focus group meetings, the following software requirements were identified and integrated into the information system

#### Functional Requirements

* The system should enable sorting of tips in a betslip.
* The information system should display the probability of a prediction/tip coming true
* The information system should display the total odds in a betslip
* A betslip that has a high likelihood of fruition should have an indicator
* The system should display tips on matches that span over different football competitions.
* Each match should have at least two tips
* The system should allow users to search for matches/fixtures.
* The system can include a subscription module to enable monetization.
* If the system includes a subscription module, then it should include an authentication module.
* The system could include a commenting module
* The information should link to social media pages to increase user viewership
* The system should collect user behavior data to improve on its weak points in delivery
* All pages in the system should be indexed by google.

#### Non-functional requirements

* The information system should be accessible at all times over the internet
* The system should be accessible over different devices
* The Interface used should scale on different screen sizes
* The information system should be able to scale resources with increasing traffic
* They system should not take not more than 3 seconds to load each page of the interface
* The information system should be installable as a progressive web app.
* The system should be secure to prevent stealing of user information
* The system should have good search engine optimization
* It should be easy to find a match in the system by searching on the google search engine

# CHAPTER FOUR: SYSTEMS DESIGN AND IMPLEMENTATION

## Architectural Design

### Model View Controller Framework

* Model: This handles the data representation, serving as an interface to the data stored in the database itself, and also allows the user to interact with the data without having to get perturbed with all the complexities of the underlying database. In django, this translates to the django Model.
* View: As the name implies, it represents what the user sees while on a browser for a web application. In django, this translates to the html template
* Controller: provides the logic to either handle presentation flow in the view or update the model’s data. This translates to the django View

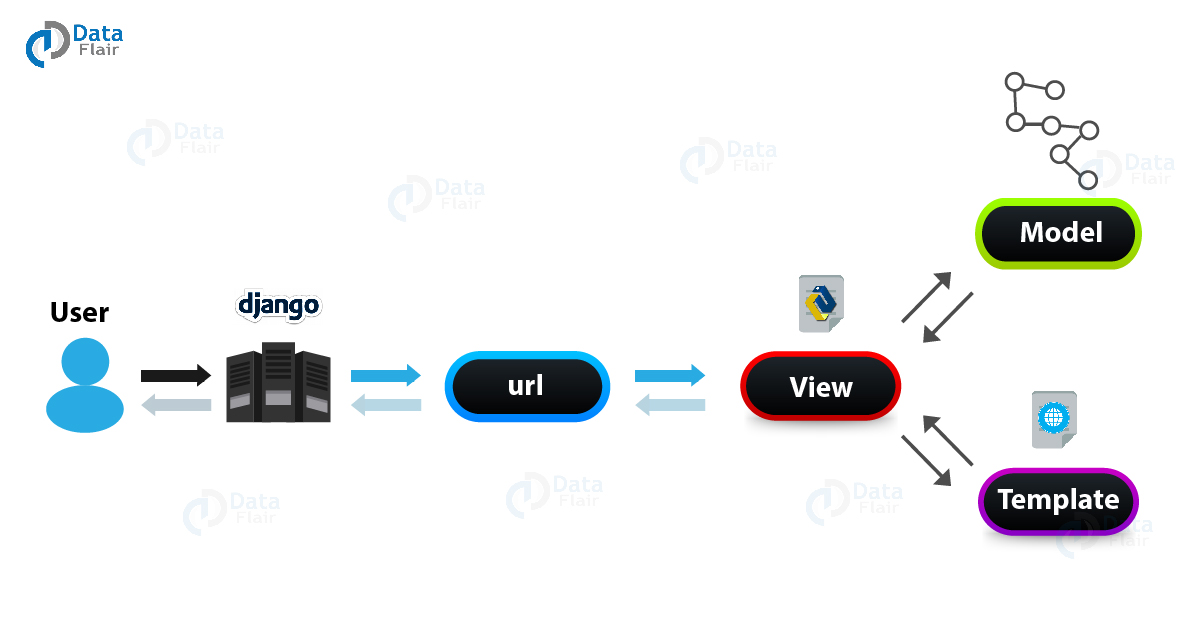


Figure 12 django MVC framework

### Cloud Deployment Architecture

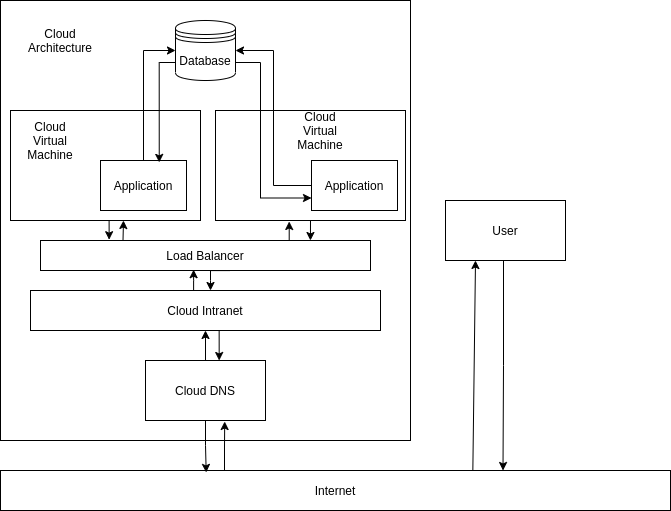


Figure 13 Cloud Deployment Architecture

### Process Flow Diagrams

* Authentication Process

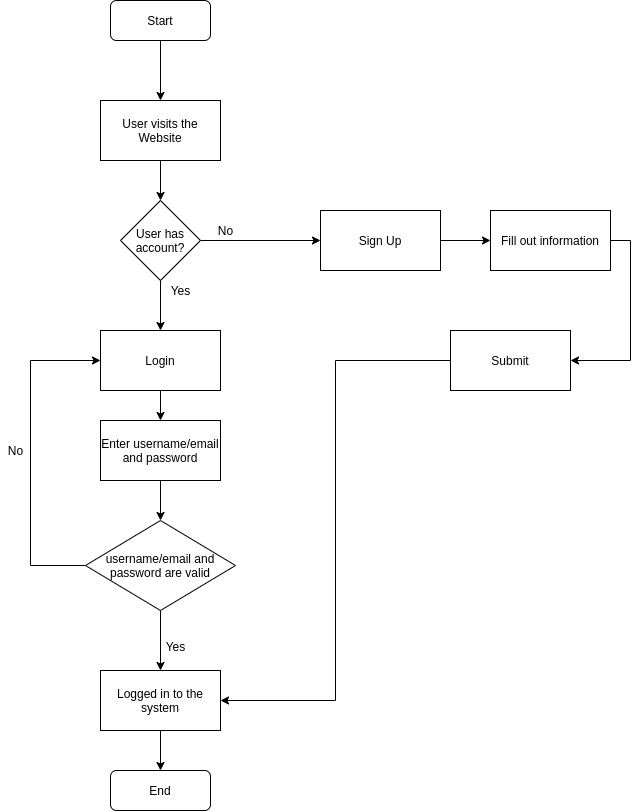


Figure 14 Authentication Process

## Interface Design

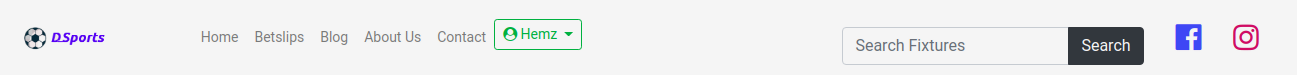
### Home Page

This page is designed to lead users to other pages. It contains a main part that shows the top betslips for a specific date and matches ordered by their competition.

This page has a search button which enables a user to search for a match in the database.

All pages in the system have a sidebar that shows the trending matches on a specific date.

In the diagram below, this page features advertisements from third parties. Social media buttons are included in the navbar to encourage users to follow the links



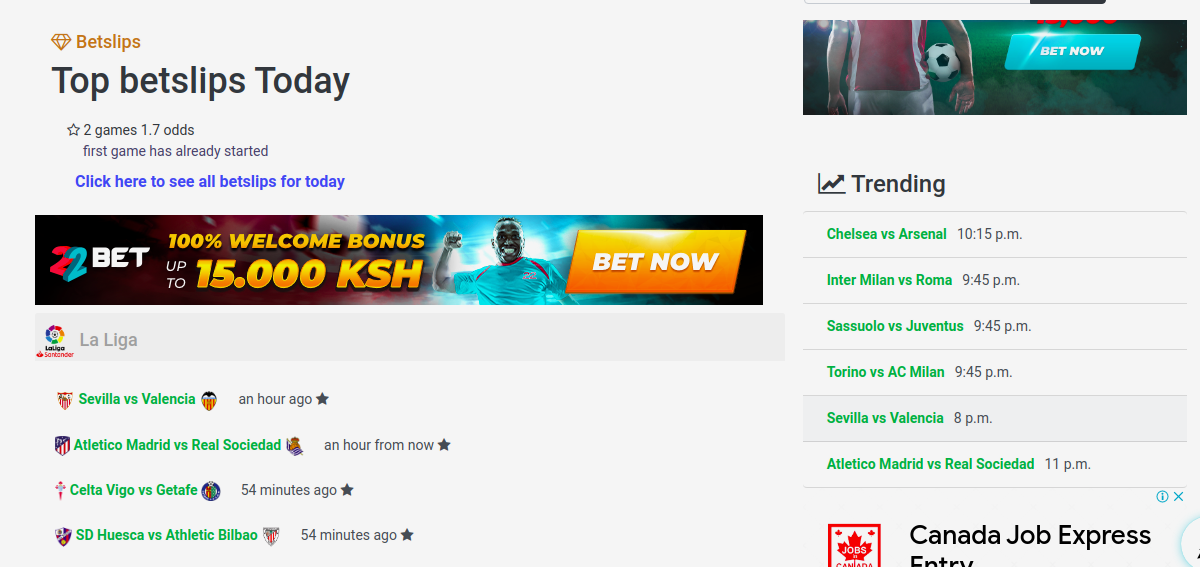
****

Figure 15 Home Page

### Single Match

This page shows the tips for a match. It shows the competition which in the diagram below is ‘eredivisie’, the time the match will start and the two teams playing against each other

In the diagram below, the match result tip shows that Ajax will win the match. The probability for Home win i.e Ajax to win, is 99.53%. The probability for a draw is 0.46% while the probability for VVV-Venlo to win is 0.01%

Other than the tips, this page also shows selected tips. The selected tips are derived using logic that selects the most probable tips to come true out of all the given tips.

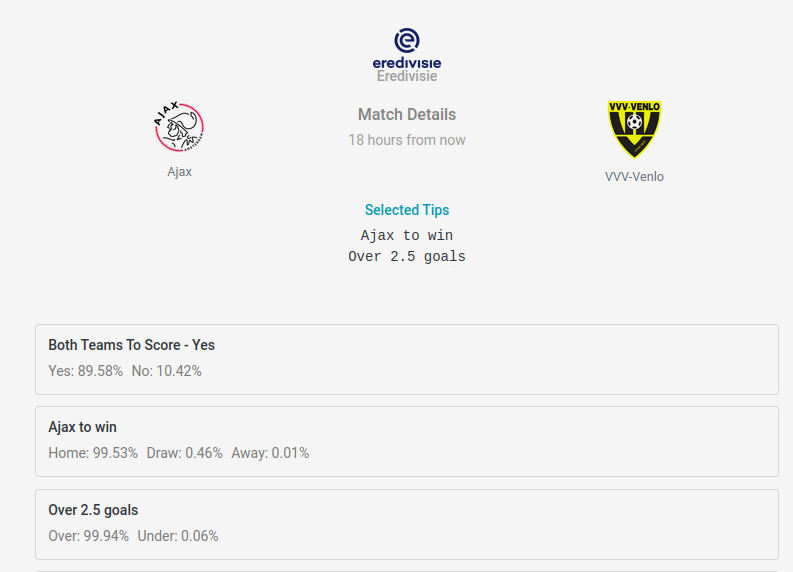


Figure 16 Single Match Page

### Betslips

This page shows a betslip. A betslip is a combination of tips. The tips in a betslip should not have the same match. For each betslip, the total odds of all the probabilities are shown.

If a betslip is marked as featured, it shows that it has a high likelihood of success. A human decides whether a betslip should be featured or not.

This page offers functionality to enable a user to fetch betslips from a previous day. This enables users to check the performance of the betslips.

In the diagram below, this page features an advertisement from a third party

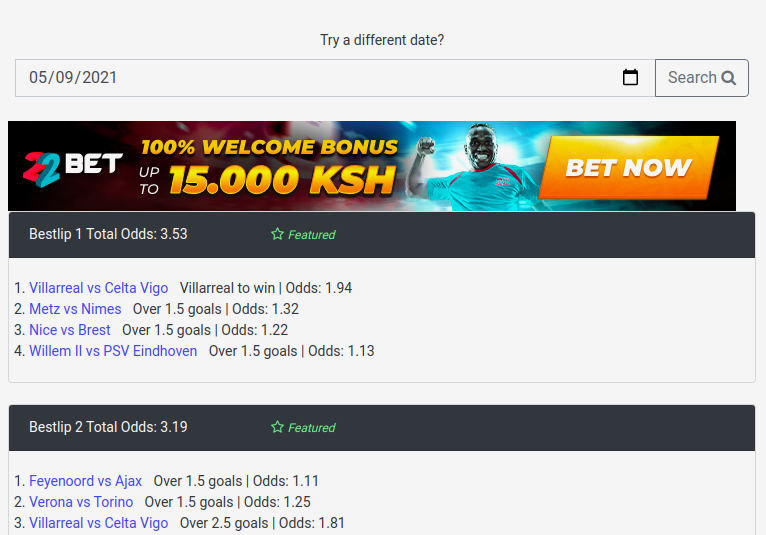
****

Figure 17 Betslips Page

### Sort Betslips

This page is limited to administrators only.

It enables an administrator to sort betslips by dragging tips from one betslip to another. On this page is where an administrator is able to mark a betslip as ‘featured’.

To disable a tip, an administrator drags it from a betslip onto the disabled tips section.

This page is designed to give the admin as much information as possible when sorting a betslip. This enables the user to create betslips according to accuracy of the tips involved and the return that a gambler would get when they place bets on all the tips in a betslip

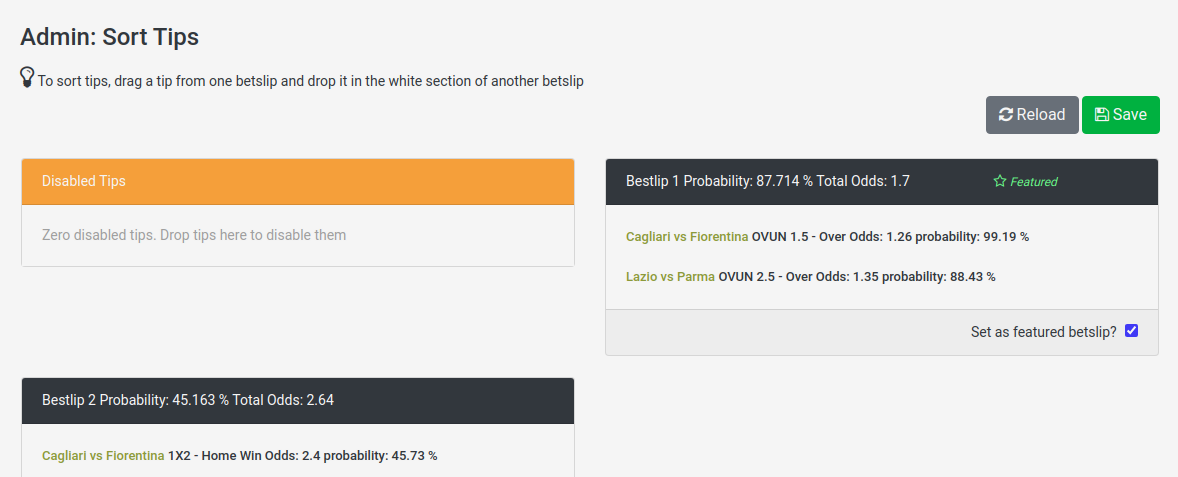
****

Figure 18 Sort Betslips Page

### Authentication

#### Login

This interface allows a returning user to log back into the site. It makes use of a captcha to identify real users from bot. The user is also able to reset their password if the forget it via email

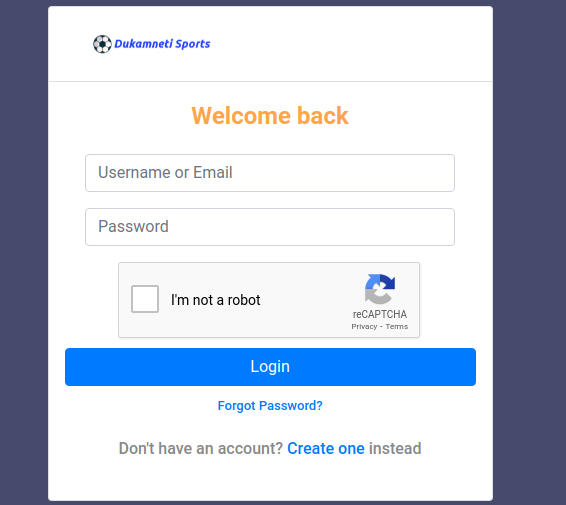


Figure 19 login page

#### Sign Up

This interface allows a new user to create a new account on the website platform.

Perks of creating an account include:

* Access to all tips when the user subscribes
* Newsletter subscription
* Personalized customer support
* Commenting and liking betting tips on the platform

#### 

Figure 20 sign up page

### Search Results

This page shows the results when a user searches for a match. The details for each match are:

1. The home team
2. The home team logo
3. The away team
4. The away team logo
5. The date of the match

The results are ordered according to date and in descending order.

This page makes use of **pagination** to limit the number of results per page to 10 results

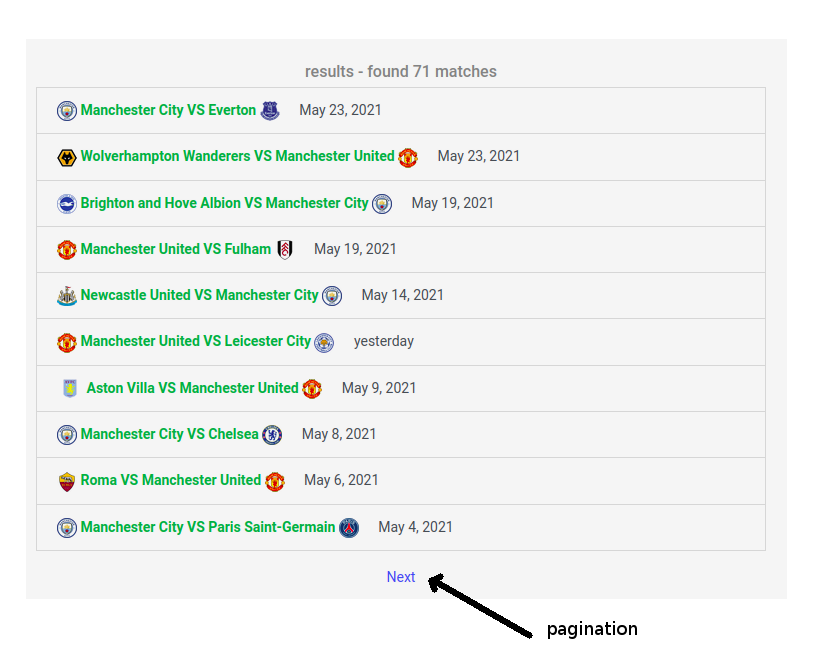


Figure 21 Search Results Page

# Database Design

### Entities

**User**

This entity represents a user of the information system. Below outlines its attributes

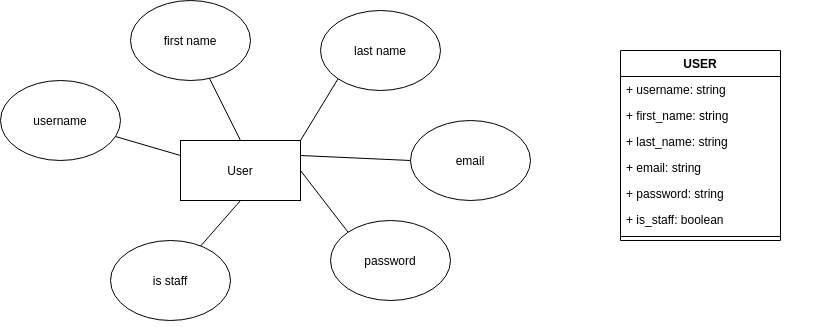


Figure 22 User Entity

**Match**

This entity represents a match in the real world

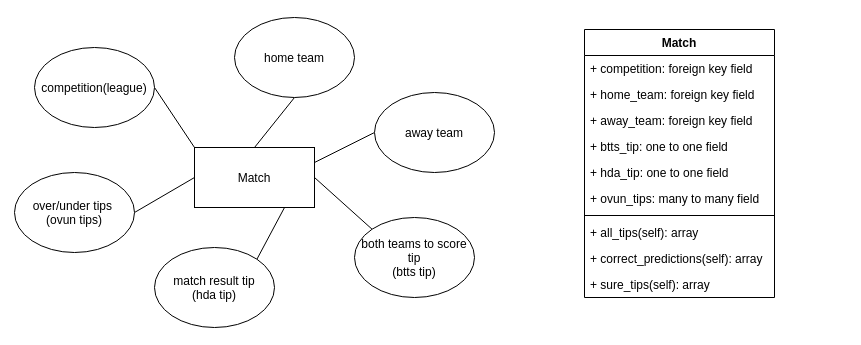
****

Figure 23 Match Entity

**Betslip**

This entity is composed of a collection of tips

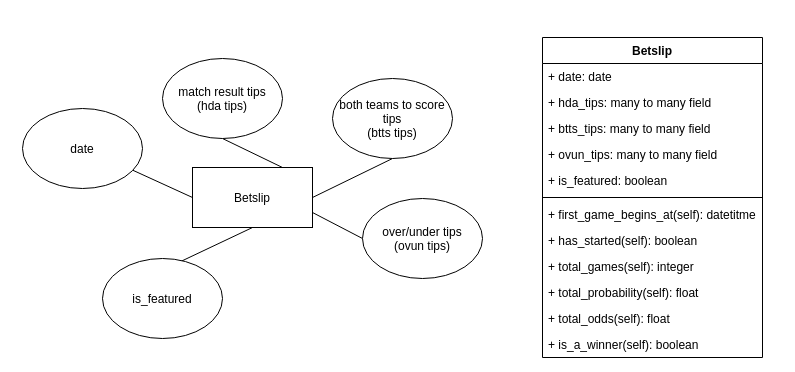
****

Figure 24 Betslip Entity

**Competition**

This entity is also referred to as ‘league’ in the real world. Examples include the English Premier and the Spanish La Liga

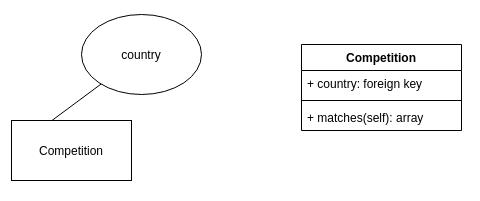
****

Figure 25 Competition Entity

**Country**

This entity represents a country in the real world

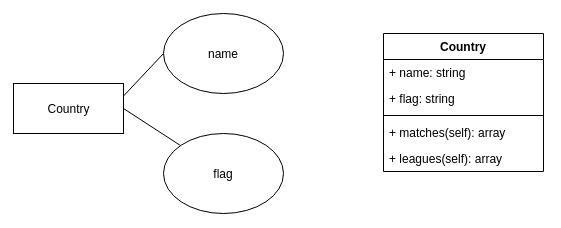
****

Figure 26 Country Entity

**Match Result Tip (HDA\_TIP)**

This entity represents a prediction of who will win the match or if the match will end as a draw

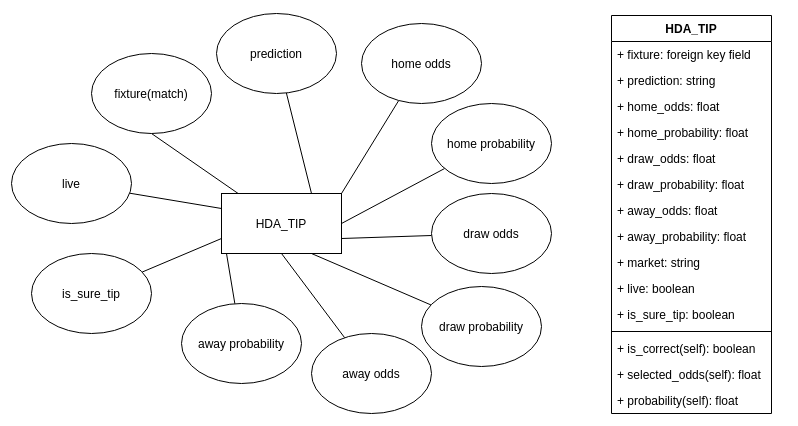
****

Figure 27 Match Result Tip Entity

**Both Teams to Score Tip (BTTS\_TIP)**

This entity represents the prediction of whether both teams in a match will score at least one goals in the fixture

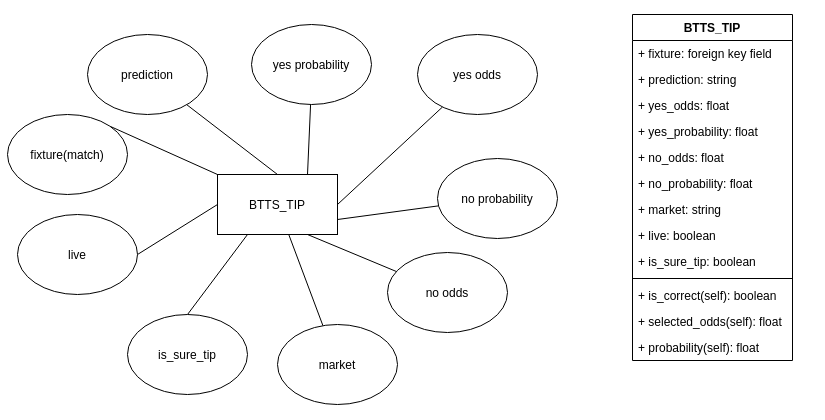
****

Figure 28 Both Teams To Score Tip Entity

**Over/Under Tip (OVUN TIP)**

This entity represents the prediction over whether the match will end in more than the specified goals, for example, whether the match will end in more than two goals.

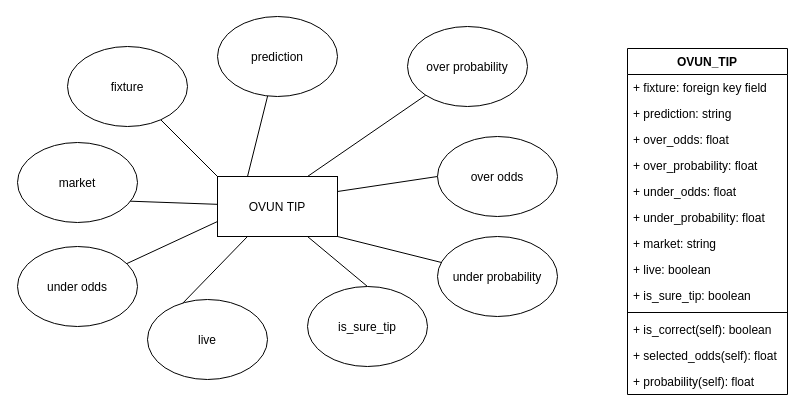
****

Figure 29 Over/Under Tip Entity

The HDA tip, BTTS tips and OVUN Tip all abstract from the class TIP as shown in the diagram below

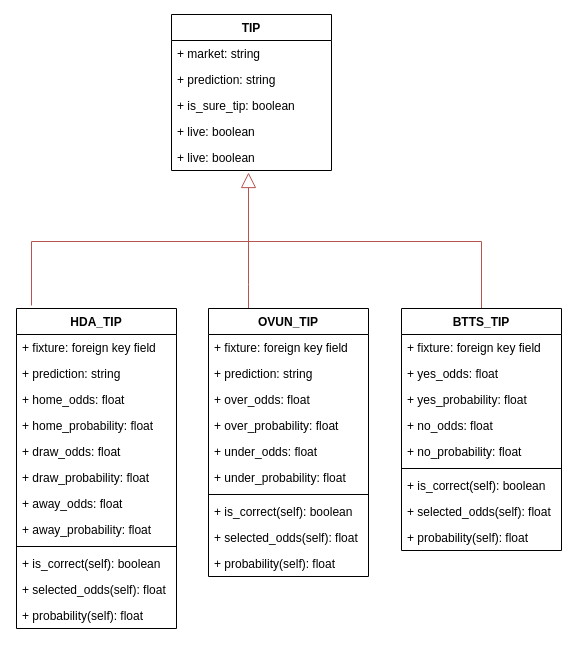
****

Figure 30 Tip Abstract Class

### Entity Relationship Diagrams

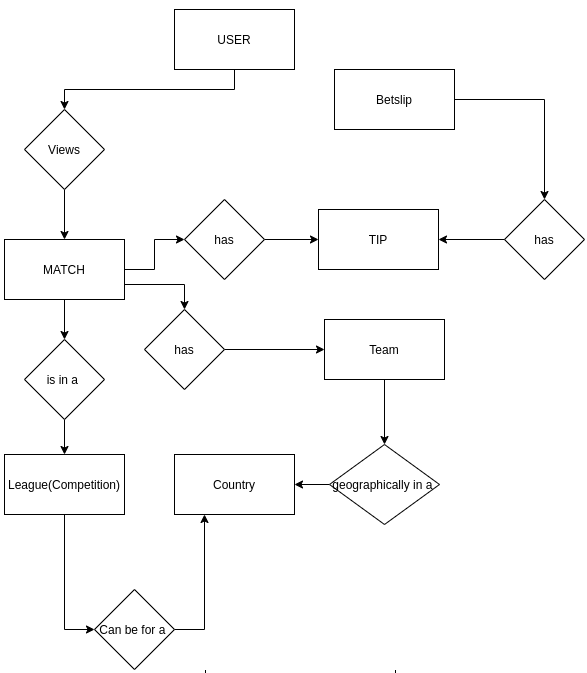
****

Figure 31 Entity Relationship Diagram

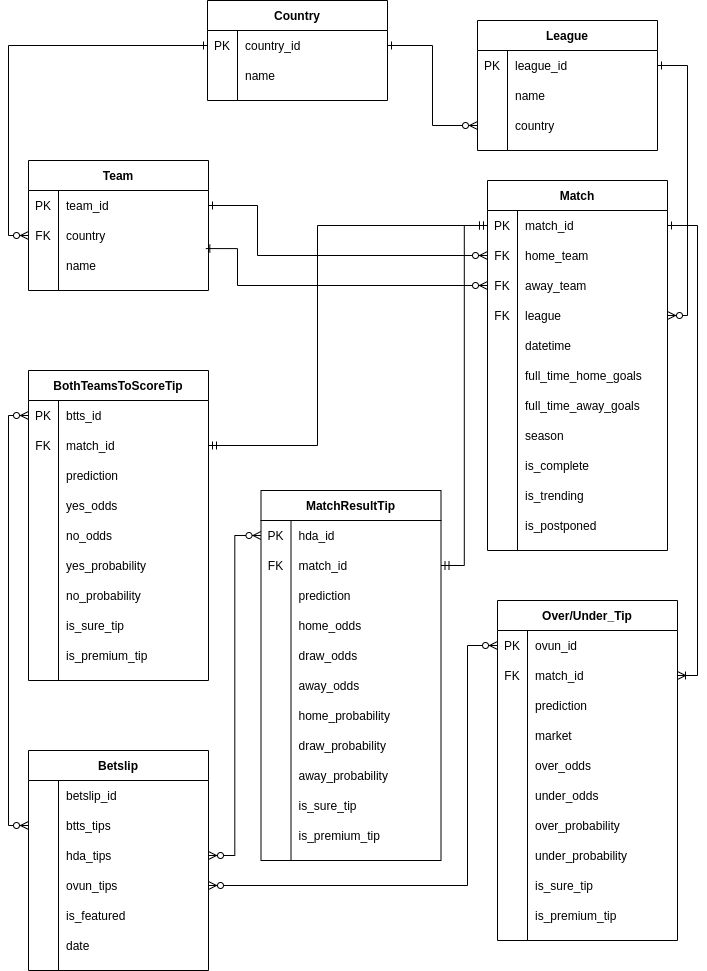


Figure 32 Database Entity Relationship Diagram

## Tools and Technologies Used

**Frameworks**

Django

Wagtail CMS

**Programming Languages**

Javascript

Python

HTML

CSS

**Databases**

Postgresql

**IDEs**

Pycharm

**Software Development Methodologies**

Prototyping

**Cloud Infrastructure**

Google Cloud

## Testing

Unit Testing

Tests were run to test individual modules of the system.

Table 3 Testing table

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Subject | Test Case | Objective | Expected Results | Actual Results |
| Authentication | User enters wrong credentials | To test rejection of authentication request | Authentication is not successful | Success.  User is not Authenticated |
| Authentication | User enters correct credentials | To test successful authentication | Authentication is successful | Success.  User is Authenticated |
| Sorting Tips | Admin sorts tips by dragging on interface | To test that tips are sorted correctly and odds are calculated automatically | Tips are sorted correctly and odds calculated automatically | Success.  Tips are sorted correctly and odds calculated automatically |
| Permissions | User tries to access a module with insufficient permissions. | To test correct redirection of user with insufficient permissions | 403 Redirect. Request to resource is forbidden | Success.  403 Redirect. Request to resource is forbidden |
| Permissions | User tries to access a module with sufficient permissions. | To test that the user is given access to the resource. | 200 OK response. User is granted access to resource. | Success.  200 OK response. User is granted access to resource |

## 

## 

## Deployment

The system is deployed on the cloud using the architecture below. The architecture allows for both horizontal and vertical scaling. Vertical scaling is possible since more virtual machines can be added to the architecture to enable load balancing. Horizontal scaling is possible since more processors and storage can be added to individual virtual machines.

A **direct changeover** is implemented since the system is entirely new. Upon system evolution, phased changeovers are to be used i.e. implementing individual modules each at a time.

The design of the system is friendly in that it requires very minimal training for target users. The training is similar to that of most social media platforms in that users are required to experiment with aspects of the system to understand them. The sports predictions system interface has few pages thus reducing the time it takes for a user to understand how it works

For the administrators of the system, training topics are:

* How to add tips to the system
* How to add and edit betslips
* How to sort tips in betslips

**Maintenance Activities**

* Monthly updates of the cloud servers
* Weekly checking of server log files to track any errors in the system
* Yearly renewal of domain
* Monthly payment of cloud hosting services
* Regular offensive and defensive security operations

# Conclusion and Recommendations

The use of a web-based betting tips information system shows significant acceptance among gamblers as a source of betting information. This is attributed to the specific attention given to the three variables in this study i.e., accuracy of information, quantity of information and ease of access of the information.

Indeed, an information system that makes use of, and improves on, these variables is bound to have acceptance in the gambling community

## Recommendations

1. The accuracy of the information on the system can be improved. This can be done by improving the machine learning models used to generate the tips.
2. Social media integration should be introduced to the system to enable the users to engage with each other and with the content in the system. This can be done by introducing commenting, liking and sharing of tips on the system
3. The system should be designed to be installable on mobile devices as an app so as to enable sending of notifications from the system to users. One way this can be achieved is by introducing ‘progressive web app’ functionality in the system.

# REFERENCES

Carl Miller, Alex Krasodomski-Jones, Josh Smith (2016) Gambling & Social Media

Hibai Lopez-Gonzalez, Alicia Rius-Buitrago, Susana Jimenez-Murcia & Mark D. Griffiths (2020) The utilization and perception of sports betting experts (‘tipsters’) among sports bettors with gambling problems: a qualitative focus group interview study, International Gambling Studies, DOI: [10.1080/14459795.2020.1844270](https://doi.org/10.1080/14459795.2020.1844270)

Susan Rice Lincolin (2009) Mastering Web 2.0: Transform your website using key website and social media tools.

BBC News. (2016). The Social Media Tipsters Who Win When You Lose. BBC News. Retrieved November 30, 2016, from <http://www.bbc.co.uk/news/blogs-trending-36580678>.

[Kalina Bontcheva](https://arxiv.org/search/cs?searchtype=author&query=Bontcheva%2C+K), [Genevieve Gorrell](https://arxiv.org/search/cs?searchtype=author&query=Gorrell%2C+G), [Bridgette Wessels](https://arxiv.org/search/cs?searchtype=author&query=Wessels%2C+B) (2013) Social Media and Information Overload: Survey Results. [arXiv:1306.0813](https://arxiv.org/abs/1306.0813) **[cs.SI]**

Freeman, B, Kelly, B, Baur, L, Chapman, K, Chapman, S, Gill, T & King, L (2014), Digital junk: food and beverage marketing on Facebook, American Journal of Public Health, 104(12), 56-64.

Griffiths, R & Casswell, S (2010), Intoxigenic digital spaces? Youth, social networking sites and

alcohol marketing, Drug and Alcohol Review, 29(5), 525.

Samantha L. Thomas, Amy Bestman, Hannah Pitt, Emily Deans, Melanie J. Randle (2015): The marketing of wagering on social media: An analysis of promotional content on YouTube, Twitter and Facebook

Feng L, Hu Y, Li B, Stanley HE, Havlin S, Braunstein LA (2015) Competing for Attention in Social Media under Information Overload Conditions. PLoS ONE 10(7): e0126090. <https://doi.org/10.1371/journal.pone.0126090>

Laura F. Bright, Susan Bardi Kleiser, Stacy Landreth Grau: Too much Facebook? An exploratory examination of social media fatigue, Computers in Human Behavior, Volume 44, 2015, Pages 148-155, ISSN 0747-5632, <https://doi.org/10.1016/j.chb.2014.11.048>. (<https://www.sciencedirect.com/science/article/pii/S0747563214006566>)

Shaoxiong Fu, Hongxiu Li, Yong Liu, Henri Pirkkalainen, Markus Salo: Social media overload, exhaustion, and use discontinuance: Examining the effects of information overload, system feature overload, and social overload, Information Processing & Management, Volume 57, Issue 6, 2020, 102307, ISSN 0306-4573, <https://doi.org/10.1016/j.ipm.2020.102307>. (https://www.sciencedirect.com/science/article/pii/S0306457320308025)

Jonathan. (2020, August 26). Tipster Service: Start Your Own Business. [*https://www.jackpotfinder.com/news/tipster-service/*](https://www.jackpotfinder.com/news/tipster-service/)

Pretorius, Arnu and Parry, Douglas A(2016): Human Decision Making and Artificial Intelligence: A Comparison in the Domain of Sports Prediction. <https://doi.org/10.1145/2987491.2987493>

Samuel, Arthur L. “Some Studies in Machine Learning Using the Game of Checkers,” IBM Journal of Research and Development 44:1.2 (1959): 210–229.

Awad M., Khanna R. (2015) Machine Learning. In: Efficient Learning Machines. Apress, Berkeley, CA. <https://doi.org/10.1007/978-1-4302-5990-9_1>

Robert J. Williams, Robert T. Wood, Jonathan Parke(2012): Routledge International Handbook of Internet Gambling

Romney, J.(1995, October): Tiny Liechtenstein offers first-ever Internet Lottery

[104th United States Congress](https://en.wikipedia.org/wiki/104th_United_States_Congress). National Gambling Impact Study Commission Final Report. <https://govinfo.library.unt.edu/ngisc/reports/fullrpt.html>

"Internet Gambling ." Gambling: What's at Stake?. . Retrieved April 15, 2021 from Encyclopedia.com: <https://www.encyclopedia.com/reference/social-sciences-magazines/internet-gambling>

“How to predict the score in football betting”, Bettingwell < [**https://www.bettingwell.com/sports-betting-guide/football-bettors-guide/how-predict-score-football-betting**](https://www.bettingwell.com/sports-betting-guide/football-bettors-guide/how-predict-score-football-betting)>

“Introduction to BigQuery”, Towards Data Science <<https://towardsdatascience.com/introduction-to-bigquery-ml-e746a3eaa28d?gi=c752839b9d2e>>