Identification of Inefficiencies in Football Betting Markets using a Statistical Approach and Artificial Intelligence Techniques

MSc Research Project
Programme Name

Donal Collins
Student ID: 21212431

School of Computing
National College of Ireland

Supervisor: Dr. Catherine Mulwa

National College of Ireland
Project Submission Sheet
School of Computing

Student Name:	Donal Collins
Student ID:	21212431
Programme:	Programme Name
Year:	2024
Module:	MSc Research Project
Supervisor:	Dr. Catherine Mulwa
Submission Due Date:	12/08/2024
Project Title:	Identification of Inefficiencies in Football Betting Markets using a Statistical Approach and Artificial Intelligence Techniques
Word Count:	XXX
Page Count:	24

I hereby certify that the information contained in this (my submission) is information pertaining to research I conducted for this project. All information other than my own contribution will be fully referenced and listed in the relevant bibliography section at the rear of the project.

ALL internet material must be referenced in the bibliography section. Students are required to use the Referencing Standard specified in the report template. To use other author's written or electronic work is illegal (plagiarism) and may result in disciplinary action.

Signature:	
Date:	11th August 2024

PLEASE READ THE FOLLOWING INSTRUCTIONS AND CHECKLIST:

Attach a completed copy of this sheet to each project (including multiple copies).	☐
Attach a Moodle submission receipt of the online project submission, to each project (including multiple copies).	☐
You must ensure that you retain a HARD COPY of the project, both for your own reference and in case a project is lost or mislaid. It is not sufficient to keep a copy on computer.	☐

Assignments that are submitted to the Programme Coordinator office must be placed into the assignment box located outside the office.

Office Use Only	
Signature:	
Date:	
Penalty Applied (if applicable):	

Identification of Inefficiencies in Football Betting Markets using a Statistical Approach and Artificial Intelligence Techniques

Donal Collins
21212431

Abstract

Football or soccer matches have long been the focus of analysis, prediction and gambling. The introduction betting exchanges has transformed betting markets into liquid and responsive markets which share characteristics with financial markets. Focusing on the English Premier League, this study develops a methodology to estimate strength or "Form" of each team prior to each match. Using these Form statistics, a measure of home advantage and implied probabilities of exchange market odds, three models have been developed to predict to outcomes of matches using a modified Poisson distribution method, Extreme Gradient Boosting and Multi-layer Perceptron machine learning methods. To test the performance of the models a simulated betting strategy was employed where model probabilities were compared to Betfair Exchange market odds of bets transacted immediately before match start times. The Extreme Gradient Boosting model generated a 17.2% return on the held-out dataset of the English Premier League 2023-24 season. All three models generated returns of greater than 10% on average when evaluated on held out English Premier League matches. To further demonstrate efficacy, the models were tested on German Bundesliga and Spanish La Liga matches where the approaches were also profitable on average, though, at a lower level. The study demonstrates that football betting markets have inefficiencies which can be identified using the methodology to measure team strength and the machine learning models described.

Keywords: Football Match Prediction, English Premier League, Betting Exchange Odds, Machine Learning Models.

1 Introduction

The efficiency of financial markets has been researched and analysed by academics, investors and professional experts for a century or more. Betting exchanges have created a market place in the world of sports betting which has a lot of similarities to modern stock markets. On sports betting exchanges, customers offer market odds as well as place the bets. Each transacted bet has a customer both sides of the bet. The exchange acts as an intermediary and collects a transaction fee but unlike traditional bookmakers do not take on any risk. In 2000 Betfair Exchange was launched and it has become the dominant player in the betting exchange market. For high level football matches like the English

Premier League markets are very liquid and responsive to new market information. Odds are updated frequently by a large customer base and large volumes of bets are placed on the platform. In summary the Betfair Exchange functions as a deep and liquid market making it an ideal data source to test market efficiencies.

1.1 Motivation and Background

There is a large volume of research in the area of financial market efficiency going back to the mid twentieth century. "The Behaviour of Stock-Market Prices" Fama (1965) was a foundational paper which introduced the "the market efficiency hypothesis". It argues that stock markets are efficient and incorporate all known information into security prices. The implications of this theory are that stock price movements follow a "random walk" pattern. Future price movements are unrelated to past movements and they cannot be predicted.

Subsequent research challenged the notion of market efficiency by revealing instances of predictability in asset prices. Bondt and Thaler (1989) found mean-reverting tendencies in asset prices and evidence of historical financial bubbles followed by and price crashes. Such findings underscore the complexities inherent in market dynamics and fuel ongoing debate regarding the efficacy of market efficiency theory in explaining real-world phenomena.

The field of economics evolved to embrace behavioural economics, integrating psychology to explore decision-making processes. Kahneman and Tversky (2013) highlights cognitive biases and emotional influences on decision-making, disrupting the traditional rationality assumption and has become widely accepted.

This study examines football betting markets in the context of market efficiency. The study will introduce a novel approach to estimating the defensive and offensive strength of football teams referred to as team "Form". It will do this by using iterative approaches to find the optimal parameters which describe the evolution of team form over time. This use of feature engineering techniques aims to address limitations of earlier studies produce data features which lead to better fitting models. The study utilises football results and statistics from the English Premier League over ten seasons in the development of the form features.

This project takes account of home field advantage and uses it as a feature in the prediction of the result of sporting events. Krieger and Davis (2022) investigated home field advantage in American football and the impact of reduced spectators in stadiums during 2020 due to the Covid 19 pandemic. The study showed a large and modestly increasing over time home field advantage in the sport from 2007 to 2019. A dramatic shift was experienced in 2020 however, when no spectators or significantly reduced numbers of spectators were allowed into stadiums to attend matches. The mean home field advantage for the 2007 to 2019 period of 7.06 points reduced to 4.62 points when spectators' numbers were reduced in due to pandemic restrictions. The advantage falls further to 2.32 points for match events with no spectators allowed in the stadium. Similar observations have been made in the football matches and therefore home field advantage is used as an input feature in the problem of prediction football match results in this study.

1.2 Research Question

Much attention over many decades has been given to the study of efficiency in financial markets. The study of sports outcome prediction has been more sparse and those studies have tended to measure prediction accuracy as the performance metric. This project proposes to build an effective model of football matches and compare the model results to the most liquid betting market in order to address the following research question:

Research Question: Can a statistical approach to estimating the strength of football teams, combined with artificial intelligence techniques reveal inefficiencies in football betting markets?
Sub Research Question: Can the methods developed, applied to football betting markets, generate a profitable betting strategy?

1.3 Research Approach and Objectives

1.3.1 The "Form" concept

Followers of sports teams and commentators use the term "form" as a vague and broad term to describe how well a team is playing at a point in time. This study uses the "form" concept as the central method of evaluating a team's strength. It formalises the concept into two metrics: "Goal Scoring Form" and "Goal Conceding Form". These metrics evaluate how many goals a team would be expected to score and concede against their league opposition on average.

Much of the effort in the study involved methods of calculating team form. Using historical football statistics form numbers using different approach and a range of parameters within the same approach. The results of the methods were evaluated to determine which approach had the most predictive power. While all the statistics used to generate the form metrics are historical, the usefulness of the numbers is in their ability to estimate future outcomes.

The approach taken in the study is intended to be a broad methodology which can be applied to any football league. Professional football leagues around the world have broadly similar structure with local nuances. Promotion and relegation ensure the composition of the league changes each year with three teams dropping out of the English Premier league and three new teams being added. The off-season between the end of one season the start of the following season can see significant changes in the makeup of individual squads as teams loose star players or strengthen their squads. Reflecting these off-season changes in the "Form" statistics are a pivotal part of the study to ensure that the inputs to models are useful.

1.3.2 The market: Betfair Exchange

This study utilises historical betting data from the Betfair Exchange over the period of the study. Betfair Exchange allows customers to offer odds on sporting events as well as placing bets. This exchange model leads to a dynamic market where the odds of an event outcome fluctuate over time incorporating new information or adjusting to market overreactions. Odds available are updated frequently by a large customer base, large volumes of bets are placed on the platform and the markets are responsive to new market information. Betfair Exchange functions as a deep and liquid market making it an ideal data source to test market efficiencies. In this study, exchange odds immediately prior to the match starting time are captured for each match event in the ten seasons. These odds are used in two ways:

- The inverse of an event outcome odds is the implied probability of the event occurring. These implied probability statistics are considered as a crowd sourced input feature in the data modelling process.

- The exchange odds captured immediately prior to considered to be the market for the purposes of efficiency testing.

1.3.3 Prediction models

Combining the form, home advantage and market input features, this study takes three approaches to predicting the football match results:

1. **Poisson Distribution:** A probability of each team scoring 0 to 6 goals is calculated using a Poisson frequency distribution. A matrix of possible scoring outcomes and the probabilities of each are then calculated. The probabilities of home win, draw and away win are then summed for each match.

2. **Traditional Machine Learning Algorithms:** Ensemble models (Random Forest, Bagging, Gradient Boosting and Extreme Gradient Boosting), K Nearest Neighbours, Gaussian Naïve Bayes and Support Vector Machine classification algorithms are trained on the input features in order to generate match outcome probabilities.

3. **Multi-layer Perceptron Algorithms:** Deep learning neural network models have been trained on the input features described in order to generate match outcome probabilities.

Model results probabilities generated by each of the three approaches be compared to Betfair Exchange match odds transacted immediately prior to the start of each match event. The predictions generated by the Extreme Gradient Boosting (XGBoost) and Multi-layer Perceptron (MLP) algorithms reveals inefficiencies in the betting markets.

The models were trained on data from 2015-16 to 2022-23 English Premier League seasons. In order to test the modelling approaches, model result probabilities were compared to the betting odds available immediately prior to the start of each match. A simulated betting strategy was applied to the 2023-24 season where bets are placed whenever available odds are significantly greater than those implied by the models. For the purposes of evaluation an arbitrary threshold of 25% was chosen.

2 Related Work in the field of Market Efficiency Theories

2.1 A Critical Review of Historical theories of Market Efficiency

In the fields of Economics and Finance "The Behaviour of Stock-Market Prices" Fama (1965) is an influential study of the efficiency of stock markets. The paper outlines evidence that stock markets are efficient and incorporate all known information into security prices. This has become known as "the market efficiency hypothesis" which is often included in modern economics and finance curricula. The hypothesis has three components.

Weak-form efficiency suggests that financial markets incorporate and reflect all publicly available information into asset prices. Semi-strong efficiency suggests that asset prices in liquid financial markets quickly and accurately adjust to new public information. Strong-form efficiency suggests that all public and private information is fully and immediately reflected into asset prices. The implications of this theory are that stock price movements follow a "random walk" pattern. Future price movements are unrelated to past movements and they cannot be predicted.

A contradictory view is presented by Bondt and Thaler (1989). This famous paper shows that equity prices do show predictable patterns. The study focuses on individual securities that have experienced large price movements and analyses them over a medium- or long-term perspective. It shows that these stocks display mean reverting tendencies. This conclusion suggests that in at least some circumstances asset prices do not follow a random walk and can be predicted. The "mean reversion theory" is therefore in direct conflict with the "the market efficiency hypothesis".

Canterbery (1999) provides more robust criticism of the market efficiency hypothesis by examining historical financial bubbles and crashes. The evidence from "tulip mania" (the 17th century bubble in the price of tulip bulbs in the Dutch Republic), the Mississippi bubble (the 18th century bubble in shares of the French government sponsored Mississippi Company) and the Wall Street Crash of 1929 showed that markets and the individuals participating in those markets could behave very irrationally. The study goes on to describe the conditions where a collective mania sets in. The "Gatsby effect" takes hold in where pure speculation takes over and individuals buy for resale rather than income.

Bondt and Thaler (1989), Canterbery (1999) and many subsequent studies provide clear evidence that "the market efficiency hypothesis" does not hold at all times. A challenge for participants in any market place and financial markets in particular is to identify when prices deviate from a rational level. Identifying overpriced or under-priced assets will allow participants to profit from mispricing or at least avoid large losses from financial bubbles. In order to explain and evaluate the presence of overpriced or under-priced assets a new branch of economics evolved.

Behavioural economics integrates psychology and economics to explore decision-making in real-world contexts departing from the assumption of rationality in traditional economic models. Kahneman and Tversky (2013) highlights cognitive biases and emotional influences on decision-making. Notable biases include loss aversion, anchoring, endowment effect, overconfidence bias, and framing effect. Daniel Kahneman's Nobel Prize in 2002, won for his work with Amos Tversky, recognised their contributions to prospect theory, which significantly impacted investing and financial market analysis, leading to widespread acceptance and further developments in the field.

2.2 Critique of Latest Methods in Modelling Sports Results

Evidence of market inefficiencies were found by Angelini et al. (2022) using Betfair Exchange pricing of English Premier League matches. As football matches have a short time frame with clearly defined informational events and results, they provide an opportunity to test Fama's theory or conversely demonstrate market biases proposed by behavioural finance proponents. This study analysed football match betting odds at match kick off and used a regression of the implied probability of the outcome. The study examines the expected values of the forecast error of different segments of the market. Outsider or unfavoured teams, defined as home teams with implied win probability of less than or equal to 24% and away teams with implied win probability of less than or equal to 14% were found to be systematically under-priced. The market placed too high a probability on the stronger team winning the match and too low a probability of the weaker team. This is called "reverse favourite–longshot bias" in betting terminology. The conclusions of the paper present evidence of inefficiencies in Betfair Exchange pricing which this project set out to find. One limitation of the study is that it analysed the markets in aggregate rather than individual match events. While its findings determined that particular segments of the market are on average underpriced, the methodology used could not evaluate individual match events. This study provides a methodology to analyse match events individually rather than by segment of the market.

A methodology to derive predictive features from football match statistics was introduced by Kozak and Głowania (2021). This approach measures the strength of each team from their position on the league table and the number of match points they have earned. Głowania et al. (2022) takes a similar approach to measuring team strength. Using league position to estimate team strength has some drawbacks however. In both studies the first five rounds of league matches are removed from that dataset due to their "relatively chaotic distribution in the league table" Kozak and Głowania (2021). While the rationale that it takes a number of matches for stronger and weaker team positions to settle makes sense the selection of five rounds is not explained and appears arbitrary. Reviews of the historical records show that position on the league table has other problematic aspects. Due to scheduling issues and other competitions, teams may not have played equal number of matches distorting league table positions at any point in time. No consideration is given to the home team advantage which is evident in the historical match statistics. Similarly, no consideration is given to the strength of the opposition so all teams are considered equal when calculating the strength metric. The methodology described in this study provides a measuring of team strength from match statistics and attempts to address the limitations apparent in the methodology described above.

Recent studies of sports prediction have achieved better results with neural network models compared to traditional machine learning models. Long Short Term Memory (LSTM) models performed slightly better than Recurrent Neural Networks (RNN) at predicting football match results in Tiwari et al. (2020). The models are designed as a classification problem with outcomes of Home Win, Draw or Away Win. Nivetha et al. (2022) adds to the team strength approach by adding "streak" features to the dataset. Streaks are defined as the number of times a team has won three matches, won five consecutive matches, lost three matches and lost five matches consecutively. This steak approach is an improvement on the previously referenced approaches but it has the start of season limitation evident in the Kozak and Głowania (2021) and Głowania et al. (2022) approach. As the teams participating each league change each season due to relegation, these statistics will not be available at the start of any season for all league teams. Nivetha et al. (2022) also models the match event as a classification problem using RNN and LSTM variations. Again, the LSTM model proves the best fit and returns the highest accuracy statistics.

Most studies in the field have focused on the machine learning aspect of the problem of predicting football match results. Baboota and Kaur (2019) put emphasis on feature engineering to ensure quality inputs to their models. They use the streak concept presented by Nivetha et al. (2022) and add a "Weighted Streak" variation. This is a similar formula for team strength with higher weighting to more recent matches. Baboota and Kaur (2019) places emphasis on home and away dynamics recognising the impact this has on results. They also add a team “Form” coefficient which measures team strength relative to opposing teams. These statics are updated after each match with the form coefficients of the outperforming team increasing and form coefficients the under performing team decreasing. As it is a relative performance measure, form coefficients of a team may increase even if it loses the match once it outperforms the expected level. This relative performance metric is a significant improvement on the measures of team strength used in the other studies referenced. Feature importance statics presented in the study confirm that the Form statistics have more predictive power than the other statistics.

The approaches described have focused on identifying trends in the historical match statistics using RNN and LSTM models. This approach will always have the limitation of the the football season cycle where each year the set of teams and makeup of squads in any league changes and the statics must reset at the start of each season. This study proposes a new approach to measuring team form which will be key to producing accurate football match predictions. It treats each match as an independent event where only the strength of the two teams and and home advantage are input to the models. While the approach takes some inspiration from the form coefficient approach introduced by Baboota and Kaur (2019) it aims to improve the method. Previous studies have either ignored or not adequately addressed the home field advantage effect. Işın (2023) looked at this attribute in Turkish football and investigated if referee bias could be identified as the driver of the effect. It concluded there was little evidence that referee bias caused the Home Advantage effect and speculated that crowd noise, unfamiliarity with the field or travel fatigue may be more important factors. Whatever the underlying cause the pattern has long been observed and is consistent across leagues though the size of the advantage has varied over time and across leagues. This study calculates a rolling average “Home Advantage” metric and uses this as a model input.

3 Research Methodology

3.1 Football Prediction Methodology

The Knowledge Database Discovery (KDD) Methodology was chosen as a relatively simple but structured and robust methodology to follow in the delivery of this project. Figure 1 below describes the steps involved.

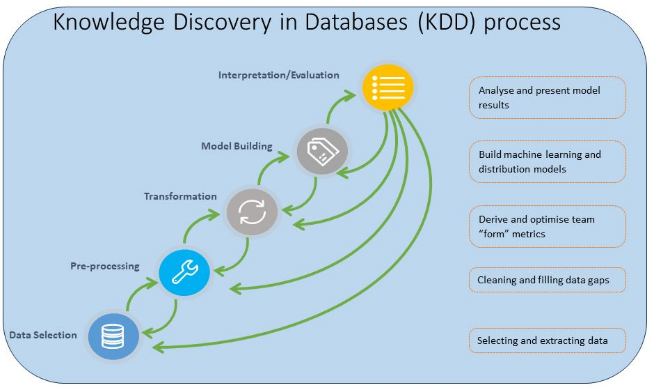

Figure 1: KDD Process Diagram

3.2 Design Process Flow

The design approach is presented below. The Betfair Exchange odds are used to imply probabilities when are then inputs to the Multi-layer Perceptron and Traditional Machine Learning algorithms. The Betfair Exchange odds are also used to measure the performance of the models.

Historical football statistics data is used to estimate team "Form". These Form measures are input features to the Multi-layer Perceptron algorithms, Traditional Machine Learning algorithms and the Poisson Distribution model algorithm.

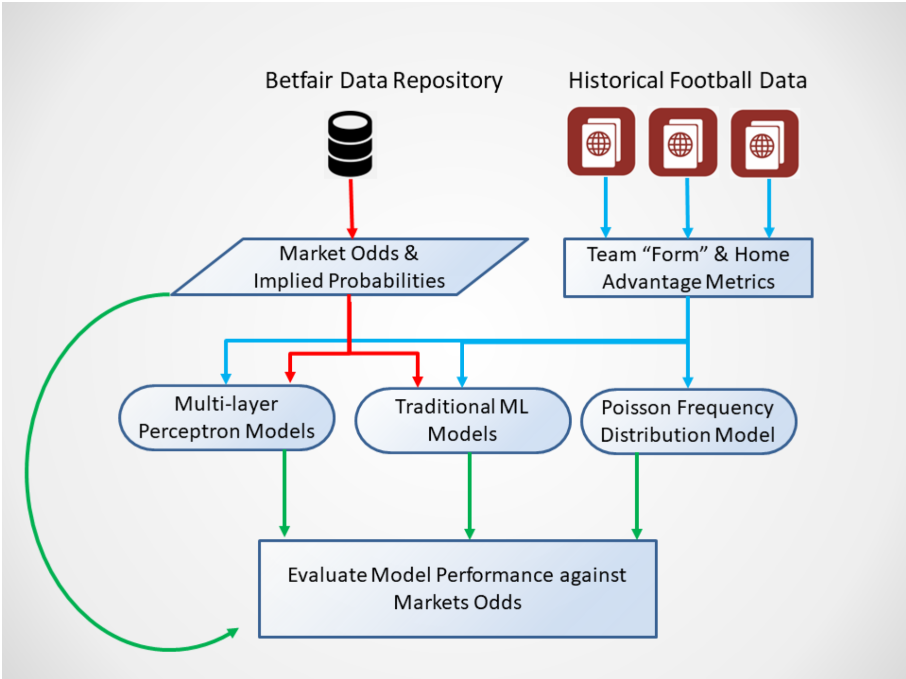

Figure 2: Design Process Flow Chart

4 Feature Engineering and Implementation

4.1 Home Advantage Metric

The Home Advantage metric is a key input in each of the three model types built in the study. Each model considers the long-observed pattern in football matches and sporting generally, that on average teams playing in their home stadium outperform teams playing away from home. This study measures the home advantage in terms of the number of home team goals scored minus number of away team goals scored.

The variation in these season averages demonstrate that the home advantage strengthens and weakens over time. Particularly striking is the 0.01 number indicating that home advantage all but disappeared during the 2020-21 season. A method of accurately estimating the advantage at any period will be important in order to build useful models.

Figure 3 shows the 380 match simple rolling average home advantage for the three seasons with the most volatility: 2019-20, 2020-21 and 2021-22. The chart also includes 190 and 380 match exponentially weighted averages of the home advantage statistic.

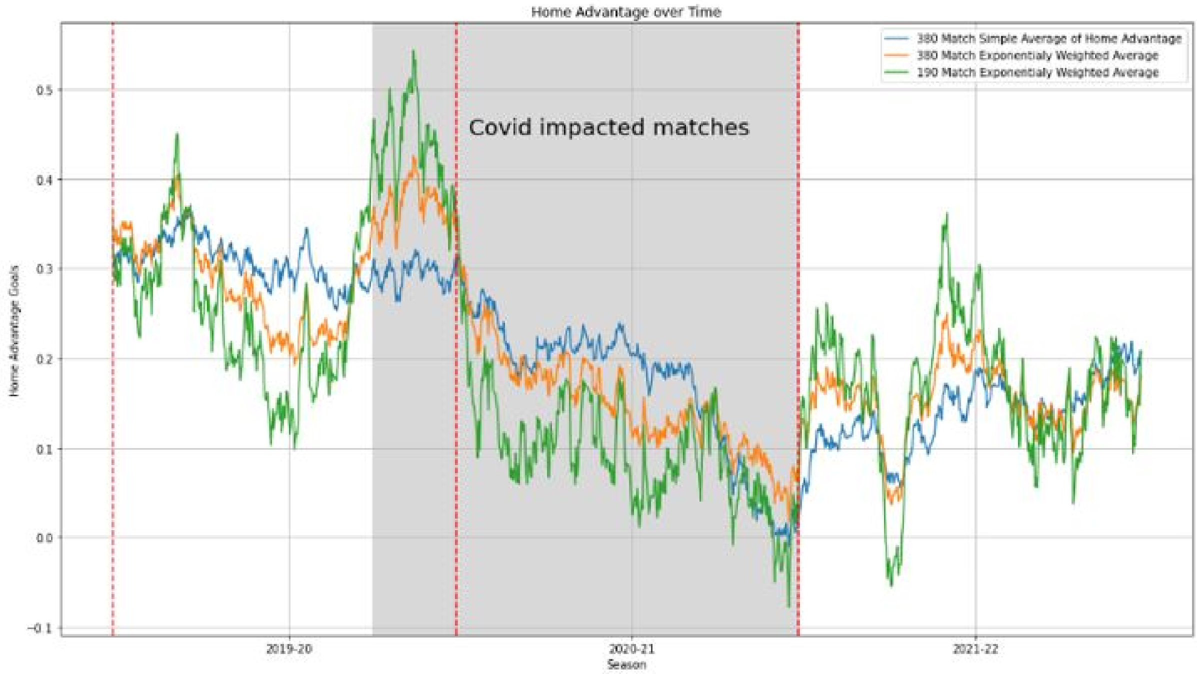

Figure 3: Rolling and Exponentially Weighted Averages

Exponentially weighted moving averages place higher weight to more recent observations and less weight to more distant observations.

The orange and green lines show the evolution of the 380 match and 190 match exponentially weighted averages. These lines are more volatile than the blue simple moving average line. As the exponentially weighted averages are more responsive to recent trends, this study expected they would more closely track home advantage and have smaller errors when compared to the observed goal differences in upcoming matches.

The 380-match simple rolling average is the best performing averaging method whether measured by Root Mean Squared Error or Mean Absolute Error. The smoothing effect of the long averaging period ensures it outperforms the shorter and more responsive methods. This rolling average will therefore be used as a model input.

4.2 Previous Season Form Statistics

At the start of any football season there are limited metrics to predict any team's performance level in the upcoming games. This study judges the previous season performance to be a good reference point to predict a team's form at the start of the following season. However, during the off-season i.e. the period between the end of the previous season and the start of the new season, teams may have a considerable change in personnel. Some new players join the squad and some players depart for other teams or retirement. There are often changes in the coaching and management staff also. Teams will usually play friendly matches in advance of the new season but the emphasis in these games are on fitness and tactical preparation rather than results so these matches were not considered a useful predictor of the team performance level. This study starts with team performance from the previous season and then adjusts for off season squad changes.

Two approaches have been taken to measure team performance for the previous season:

- An averaging approach where goals scored and goals conceded as well as expected goals for and expected goals against are averaged for each team over the season.
- An iterative optimisation approach where the goal score and goal concede form of each team is adjusted so that the difference between projected scoreline and actual scoreline is minimised.

The effectiveness of each method is evaluated by how closely future scorelines are predicted by the form numbers generated. Evaluating the difference between projected and observed scorelines allows the methods to be compared. In addition, it allows the parameter values to be tested. The averaging approach is simpler to implement and requires less processing. The iterative method presented below produced lower errors when compared against future scorelines and was therefore chosen as form measure to be used in subsequent steps of the project.

4.2.1 Iterative approach to measure team performance

The first step in the iterative approach is to measure the average goal scoring form and goal conceding form of each team over the season. This is computed by applying the formulas below to each team in the league:

$$avgGoalsFor = \frac{\sum_1^n GoalsScored_n}{n} \tag{1}$$

$$avgGoalsAgainst = \frac{\sum_1^n GoalsConceded_n}{n} \tag{2}$$

where:

- n is the number of matches played during the season
- GoalsScored and GoalsConcede are the number of goals scored and goals conceded by the team being evaluated in a particular match

The next step is to take the average team form statistics and calculate how closely the results of each match align. This is done by assigning the calculated form to each team at the start of the season. For each match in the season the projected scoreline is calculated using those form numbers. A “delta” value representing the difference between projected and actual performance is measure for each match. The form of the competing teams is then adjusted up or down to reflect their over or under performance. This process is repeated for the entire season of 380 matches. When viewed over the entire season the “delta” values provide a measure of how closely the form statistics predict a team’s performance. The aim of the optimisation approach is to minimise delta values in order to find the form values that best reflect each teams’ level. The formulas below show the calculations performed on each match over the season:

1. Projected goals formula for each match

$$projHomeGoals = (HTGSForm + ATGCForm + HomeAdv)/2 \tag{3}$$

$$projAwayGoals = (ATGSForm + HTGCForm - HomeAdv)/2 \tag{4}$$

2. Goals delta for each match

$$deltaHomeGoals = (xGHome * xGWeight + homeGoals * (1 - xGWeight)) \\ - projHomeGoals \quad (5)$$

$$deltaAwayGoals = (xGAway * xGWeight + awayGoals * (1 - xGWeight)) \\ - projAwayGoals \quad (6)$$

3. Updated form statistics for each match

$$postHTGSForm = preHTGSForm + deltaHomeGoals * formAjustProp \quad (7)$$

$$postHTGCForm = preHTGCForm + deltaAwayGoals * formAjustProp \quad (8)$$

$$postATGSForm = preATGSForm + deltaAwayGoals * formAjustProp \quad (9)$$

$$postATGCForm = preATGCForm + deltaHomeGoals * formAjustProp \quad (10)$$

The next step is to adjust the starting form of a team up and down and run the season simulation. Summing the delta values over the season provide a measure of fit. If the delta value has increased the new form number is discarded. If the delta reduces the new form number provides a better estimate of the team's performance level and the form statistic is updated.

This process is repeated for each of the twenty teams in the league. The whole cycle is then repeated until no further adjustments can be made to reduce the delta. In the implemented algorithm, few adjustments were observed after three iterations. Therefore, for each set of parameters and for each season the iterative process was run for three cycles and the minimum delta value recorded.

4.3 Off season adjustments to determine start season "Form"

Historical analysis shows team performance is highly correlated to the squad value Gerhards and Mutz (2017). Transfer markets are closed for most of the season meaning that football clubs are unable to buy or sell players. Transfer markets are opened during the off season and clubs usually have significant turnover of players as some players retire, players are sold and new players are purchased. The "Off season adjustments" captures the increase or decrease in each teams' squad values in the off season between the end of one season and the start of the following season.

Three teams are relegated from the English Premier League each year and three new teams promoted from the second-tier league known as the Championship. Performance metrics for the promoted teams, where they have faced lower-level opposition, are not comparable to Premier League performance metrics. This study uses the squad value to estimate the start season form of the three promoted teams.

Ordinary least squares regression lines are calculated for each season to provide a measure of the performance level one would expect at a given squad value.

4.3.1 Form adjustment to reflect squad turnover during off-season

For the 17 teams which remain in the English Premier League from the end of one season to the start of the following starting form is calculated as:

$$\begin{aligned} GSForm \Leftarrow regGSForm * squadValuePassThrough \\ + prevGSForm * (1 - squadValuePassThrough) \end{aligned} \quad (11)$$

$$\begin{aligned} GCForm \Leftarrow regGCForm * squadValuePassThrough \\ + prevGCForm * (1 - squadValuePassThrough) \end{aligned} \quad (12)$$

where:

- GSForm and GCForm are start of season goal scoring and goal conceding form for a team
- regGSForm and regGCForm are estimates of goal scoring and goal conceding form based on the team's squad value at the start of the new season using the regression method described in section 4.3.
- prevGSForm and prevGCForm are the end of previous season goal scoring and goal conceding form metrics calculated as described in section 4.2.
- squadValuePassThrough is the proportion of form captured from the regression method with the remaining (1-squadValuePassThrough) captured from the previous season method

4.3.2 Form estimates for promoted teams

For the three new teams promoted to the English Premier League there are no comparable previous season statistics available and their form is estimated by identifying where their squad value sits on the regression lines.

$$GSForm \Leftarrow regGSForm \quad (13)$$

$$GCForm \Leftarrow regGCForm \quad (14)$$

4.3.3 Optimising Off Season Adjustment method

In order to find the optimal squadValuePassThrough parameter value a range of values were tested. Having collected end of previous season form statistics as described in section 4.2 and applying the Off Season Adjustment, a set of start season Form statistics are prepared for varying squadValuePassThrough values. The optimal squadValuePassThrough value will is identified as the value which minimises the difference between projected and observed team performance.

4.4 Evolving Form Statistics over Season

Using the methods described in sections 4.2 and 4.3 start season form statistics were calculated for each team in the league. As the season progresses some teams will exceed and some teams will fall behind their start season metrics. This study takes the approach of calculating a projected scoreline for each match using prior form metrics for the competing teams. Delta goals values are calculated for each match by comparing actual performance against the projected scoreline. The Delta values indicate if a team under or over performs their form metrics. By definition if one team over-performs the projected scoreline then the opposition under-performed the projected scoreline. A proportion of the over or under performance is passed to the form metrics for each team after the match. This method of updating adjusting form statistics is described in section 4.2.1. The process is repeated for each of the 380 matches in each season.

Sections 4.2 and 4.3 have described a methodology to calculate form metrics for each team at the start of a season. The method has left the choice of previous season form approach and tuning of three parameters open:

- **Previous season form:** averaging or iterative optimisation approach
- **xGWeightProp:** how much weight to assign to the expected goals performance statistic with the remining weight assigned to actual goals scored and conceded
- **formAdjProp:** how much of a team's over or under performance should be passed on to the team's form metrics for subsequent matches
- **squadValuePassThrough:** the proportion of form captured from the regression method with the remaining proportion captured from the previous season method

The study has taken the approach of testing a range of values for the three parameters on the averaged and optimised previous season measures of form. The effectiveness of each approach is measured by minimising the delta approach described in section 4.2. The parameter ranges considered are presented on figure 4.

xGWeight										
0.000	0.125	0.250	0.375	0.500	0.625	0.750	0.875	1.000		
actual goals only ignoring xG statistics						xG statistics only ignoring actual goals				
formAdjProp										
0.00	0.02	0.04	0.06	0.08	0.10	0.12	0.14	0.16	0.18	0.20
static form with no adjustment								update form by 20% of over/under performance		
squadValuePassThrough										
0.000	0.250	0.500	0.750	1.000						
previous season form only		squad value regression only								

Figure 4: Range of parameters considered in the estimate of team form

The minimum errors are found at the 0.08 Form Adjustment point on 0.75 Squad Value Pass Through line. These best fitting parameters to derive the Form metrics which will be used in the modelling process are presented below:

- **Performance measuring approach: Optimisation**
- **Expected Goals Weighting: 0.875**
- **Squad Value Pass Through: 0.75**
- **Form Adjustment Proportion: 0.08**

4.5 Betfair Market Odds and Implied Probabilities

Betfair Exchange provides historical betting transaction data on various markets it covers including football markets. The data is provided as a large volume of zipped binary files. More than 40GB of football betting transaction files were downloaded to a hard drive for processing. The files are tagged by sport and country but not competition. Python scripts were written to read the files and select the relevant information. The methodology followed the steps:

- Loop through the directories read the meta data of each file. Save the file location, Event Identifiers, Country Code and descriptive information to a flat file.
- Loop through the events listed in the flat file. For each event with Country Code = 'GB' open the associated binary file and extract all betting trades recorded. Save results to a flat file.
- Load all the betting transaction information to a Pandas Dataframe object. For each match event select the home team, draw and away team transacted bet immediately prior to the match kick off time.
- Select all match events from the Dataframe between the start and end of the season dates. From this season subset, select all events where both the home and away teams are English Premier League teams in the relevant season.
- From the season selection exclude match events which were not Premier League matches. Teams will occasionally play each other in cup competitions.
- Append missing match event data to ensure a full dataset is generated for each season. Save the results to a CSV file for each season.

Incorrect match kick-off times causes some data quality issues in the process described above. Football match schedules are often re-arranged to fit television schedules or for many other reasons. On reviewing the results of initial runs of the process, it became apparent that the times presented on Betfair's data files were not always correct. In cases where the match start time on the files was after the actual match event, the last bet transactions were no longer from before the match has started. This returned very large odds for one team and very small odds for the opposition in the cases where goals had been scored. In order to correct this, the match times were analysed and a step was added at point 3 to correct match times.

In cases where match events are missing from the Betfair dataset equivalent betting information was captured from the from Football-data.co.uk. This source provides a selection of pre match odds sourced from traditional bookmakers. While not a liquid market the odds provided by traditional bookmakers are comparable and very similar to the Betfair numbers.

The end result of process described is a dataset of odds of match outcomes captured for all matches in the seasons reviewed. These odds are the market price of each outcome. The sum of the inverse of the match odds totals to close to 100% for each match but deviation from 100% can occur if the market is not sufficiently liquid. Probabilities are normalised to account for these slight deviations.

Figure 5 two examples of odds extracted from Betfair Exchange odds and the calculated implied result probabilities. Immediately prior to the kick off of the Manchester

Match Date	Season	Home Team	Away Team	Home Odds	Home Odds	Away Odds	Home Win Probability	Draw Probability	Away Win Probability
2015-08-08 15:45:00	2015	Manchester United	Tottenham	1.72	4.0	6.0	0.581395	0.250000	0.166667
2015-08-08 18:00:00	2015	Bournemouth	Aston Villa	1.95	3.8	4.4	0.512821	0.263158	0.227273

Figure 5: Betfair Exchange odds and implied probabilities

United versus Tottenham match on 8th August 2015 bets were placed on the Betfair Exchange on all three possible match outcomes. Manchester United were backed to win at 1.72, a draw at 4.00 and Tottenham to win at 6.00. The inverse of the transacted odds is used in this study as the implied outcome probability:

- Odds Implied Home Win Probability = 1/1.72 = 0.581395 = 58.14%
- Odds Implied Draw Probability = 1/4 = 0.25 = 25%
- Odds Implied Away Win Probability = 1/6 = 0.166667 = 16.67%
- Total Probability = 58.14% + 25% + 16.67% = 99.8%

As the Total Probability does not sum to 100% the values are normalised:

- Normalised Home Win Probability = 58.14% / 99.8% = 58.25%
- Normalised Draw Probability = 25% / 99.8% = 25.05%
- Normalised Away Win Probability = 16.67% / 99.8% = 16.70%
- Total Normalised Probability = 58.25% + 25.05% + 16.70% = 100.00%

5 Implementation of Football Prediction Models

At the completion of section 4 Form, home advantage and Betfair odds implied probabilities have been generated for all matches in the dataset. The dataset has 8 features to be used in the prediction modelling process. Figure 6 presents the input features for two matches shown in the previous section.

Home Goal Score Form	Home Goal Concede Form	Away Goal Score Form	Away Goal Concede Form	Home Advantage Goals	Home Win Probability	Draw Probability	Away Win Probability
1.626792	0.987056	1.432277	1.318275	0.386842	0.581395	0.250000	0.166667
0.985031	1.457486	0.983299	1.357214	0.392105	0.512821	0.263158	0.227273

Figure 6: Input features to models

A dependent, "MatchResult" variable was added to the dataset and the variables were used to create three types of prediction models: Multi-layer Perceptron Algorithm, Traditional Machine Learning Algorithms (Random Forest, Bagging, Gradient Boosting, Extreme Gradient Boosting, K Nearest Neighbours, Gaussian Naïve Bayes and Support Vector Machine) classification algorithms and a Poisson Distribution model.

5.1 Poisson Distribution Model

A Poisson distribution is a discrete distribution that estimates the probability of an event happening a certain number of times within a given interval of time or space. The distribution is described in 15 and has one parameter, λ (lambda), which is the mean number of events. This distribution curve approach allows the Form metrics be transformed into match result probabilities in a straight forward manner. The distribution approach to match prediction can be considered a technique to measure how well the Form metrics predict match outcome without employing machine learning techniques. In effect, the distribution approach measures the quality of the input data to be used in the machine learning models.

The usefulness of Poisson distributions in football match predictions is evident in figure 7. The plot shows the frequency of the number of goals scored by the home team in the English Premier League from seasons 2015-16 to 2023-24. The chart presents a vertical line showing the average number of goals and the red line is the Poisson distribution frequency where λ is set to the average number of goals. The Poisson frequencies align reasonably closely to the observed numbers.

For the away goals curve in particular, the historical statistics deviate from the Poisson distribution. The frequency of 0 away goals is understated by the curve and the frequency of 1 away goal is overstated. To reconcile the two, an adjustment ratio has been calculated for each curve as observed historical frequency divided by Poisson predicted frequency. The Adjusted distribution is presented as the green line on the two plots.

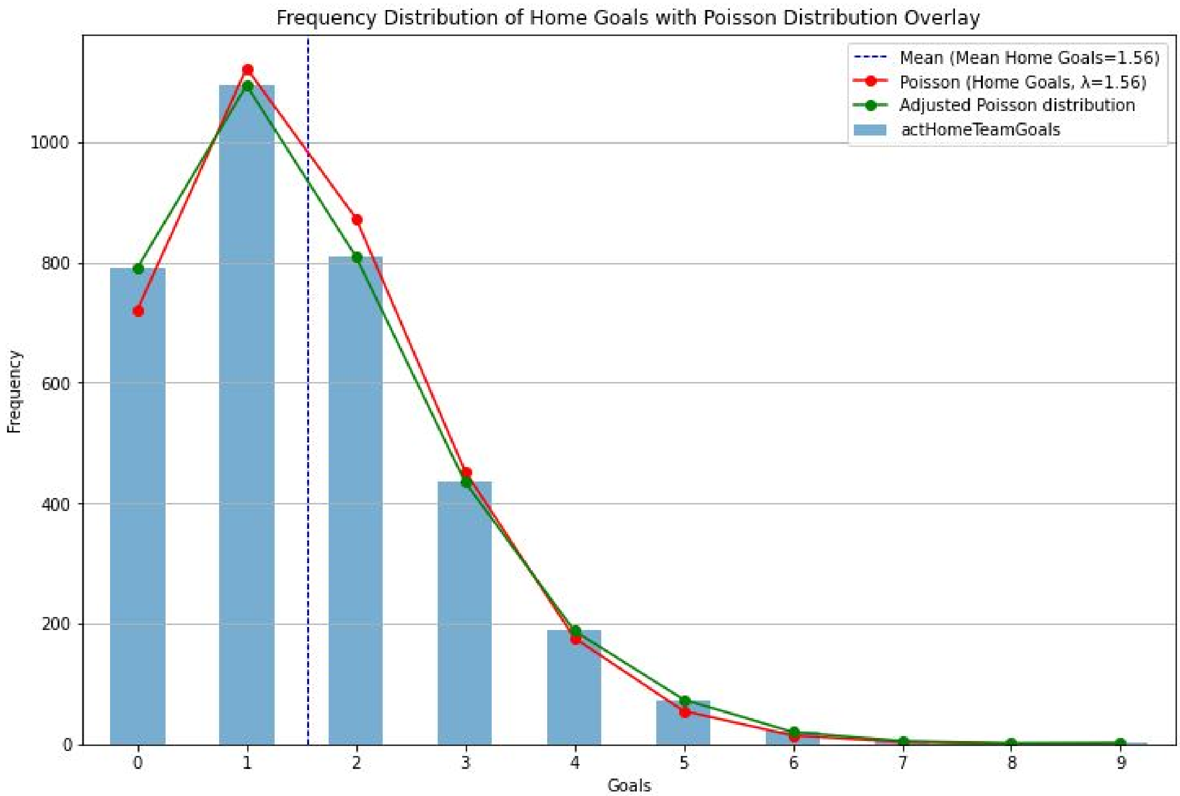

Figure 7: Home Goals Frequency with Poisson and Adjusted Poisson Distributions

The Form statistics along with the home advantage statistics are used to determine the values to input to the distribution equations. For each match event two distributions are generated: a home goals and away goals distribution. λ values for home and away goals respectively are calculated as follows:

$$\lambda \Leftarrow (HTGSForm + ATGCForm + HomeAdv)/2 \tag{15}$$

$$\lambda \Leftarrow (HTGCForm + ATGSForm - HomeAdv)/2 \tag{16}$$

Using these values of λ, the probability of each the home team and away team scoring 0 to 6 goals was calculated using the formula:

$$P(x) = \frac{e^{-\lambda}\lambda^x}{x!} \tag{17}$$

where:

- x is the number of goals
- e is Euler's number
- x! is x factorial

Figure 8 presents values for the Manchester United versus Tottenham fixture referenced earlier. The percentages represent the probabilities of the teams scoring 0 to 6 goals.

Team	0 Goals	1 Goal	2 Goals	3 Goals	4 Goals	5 Goals	6 Goals
Manchester United	19.01%	29.60%	24.68%	14.93%	7.26%	3.22%	0.98%
Tottenham	43.74%	33.86%	15.70%	5.45%	1.54%	0.34%	0.07%

Figure 8: Poisson distribution goal probabilities

Individual team goal scoring probabilities can be converted into result probabilities by calculating the product of each score result. For example, Manchester United have a 19.01% chance of scoring 0 goals and Tottenham have a 43.74% chance of scoring 0 goals. The probability of a 0-0 result therefore is 19.01% x 43.74% = 8.31%. These calculations are performed for all scoreline and the results are presented in figure 9.

Away \ Home	0 Goals	1 Goal	2 Goals	3 Goals	4 Goals	5 Goals	6 Goals
0 Goals	8.31%	12.95%	10.80%	6.53%	3.18%	1.41%	0.43%
1 Goal	6.44%	10.02%	8.36%	5.05%	2.46%	1.09%	0.33%
2 Goals	2.98%	4.65%	3.88%	2.34%	1.14%	0.51%	0.15%
3 Goals	1.04%	1.61%	1.35%	0.81%	0.40%	0.18%	0.05%
4 Goals	0.29%	0.45%	0.38%	0.23%	0.11%	0.05%	0.02%
5 Goals	0.07%	0.10%	0.08%	0.05%	0.02%	0.01%	0.00%
6 Goals	0.01%	0.02%	0.02%	0.01%	0.00%	0.00%	0.00%

Figure 9: Score probability matrix

The blue shaded area are the scorelines where the home team, Manchester United in this case, score more goals than the away team and win the match. The total home win probability in this case is 57%. The total away Tottenham probability is the sum of the green shaded values and comes to 20%. The remaining, orange shaded values represent the draw outcome probabilities and sum to 23%.

Using this methodology Poisson distribution probabilities were calculated for each match in dataset from the 2015-16 to 2023-24. The performance of the modelling method was tested using a simulated betting strategy where bets were placed wherever the available odds were 25% more profitable than the model probabilities estimated as fair value. The profitability or otherwise of the approach was measured.

5.2 Traditional machine learning algorithms

A number of ensemble learning, instance-based, probabilistic and kernel-based machine learning algorithms were trained on the 2015-16 to 2022-23 football seasons data. The 380 matches of the 2023-24 English Premier League season were held out from the training process and used to evaluate the models.

The models designed and tested are presented in table 1.

Table 1: Traditional machine learning algorithms

Algorithm	**Type**
Random Forest	Ensemble Algorithm
Bootstrap Aggregating (Bagging)	Ensemble Algorithm
Extreme Gradient Boosting (XGBoost)	Ensemble Algorithm
K Nearest Neighbours (KNN)	Instance-based Learning Algorithm
Gaussian Naïve Bayes (GNB)	Probabilistic Learning Algorithm
Support Vector Machine (SVM)	Kernel-based Algorithm

A Gridsearch type approach using a wide array of parameters was applied to each of algorithm types initially. The Random Forest, Gaussian Naïve Bayes and K Nearest Neighbours models produced somewhat promising results in initial tests but these were surpassed by Extreme Gradient Boosting (XGBoost) models.

Focus was applied to tuning the XGBoost model in order to create the best performing model possible. The following parameters we tuned:

- Number of Estimators
- Learning Rate
- Maximum Depth
- Minimum Child Weight
- Alpha
- Gamma
- Column Sample by Tree
- Reg Alpha

5.2.1 Model tuning and selection

The model performance was evaluated by generating probabilities for each match in the held out 2023-24 season. As the focus of the study is to determine if market inefficiencies could be identified, the home win, draw and away win probabilities for each match in the season were compared to the match odds available prior to the match start times. A simulated betting strategy was applied where bets were placed wherever the available odds were 25% more profitable than the model probabilities estimated as fair value. The profitability or otherwise of the model was considered the most important metric.

5.3 Multi-layer Perceptron Algorithm

Data from the 2015-16 to 2022-23 football seasons are loaded to a Python environment for model building. The 2023-24 season was held back as test data. As the historical data has been presented as independent match events, a multi-layer perceptron algorithm is the most appropriate choice of neural network model. Many design decisions were made in the process of configuring and refining the model.

5.3.1 Unbalanced Data

A consistent characteristic of football matches is the skewed nature of the results. Over a season home teams win more frequently than away teams and draw results are less frequent again. Analysis of the 2015-16 to 2022-23 English Premier League results showed 1,361 (44.8%) home team wins, 711 (23.4%) draws and 968 (31.8%) away team wins. Machine learning algorithms are sensitive to unbalanced datasets and steps were taken to ensure model predictions were appropriately balanced.

One technique used to ensure balanced predictions are generated is to use resampling techniques. In this study the Synthetic Minority Oversampling Technique or SMOTE was used. The SMOTE technique generates new records of the minority classes.

5.3.2 Choice of Loss Metric

A range of loss metrics were tested during the development of the neural network model. At one end of the scale using classification accuracy generated model configuration which performed well in terms of predicting the correct outcome classification. However, when the model probabilities, generated using the accuracy metric were considered, a tendency to over fit the data was evident. Quickly in the training process the model converged on an outcome and generated very high probabilities of that results with conversely very low probabilities of the other two result outcomes.

Metrics like Area Under the Curve performed better than accuracy and was less prone to over-fitting. A custom designed Rank Probability Score (RPS) metric was the best performing loss metric used. RPS considers how close model forecast probabilities are to observed results in terms of location (home win, draw or away win) and spread (50% probability, 60% probability, 70% probability etc.) of the forecast distribution. A lower RPS score, which indicates a better fitting model, will be calculated where a model gives the highest probability of a home win when the result is a draw relative to the when the result is an away win. A higher RPS score, which indicates a poorer fitting model, will be calculated where a model gives an 80% probability of a result that does not occur relative to a model that gives a 70% probability of the same result. The RPS formula is outlined in equation 18.

$$RPS = 1/(r-1)\sum_{i=1}^{r}(\sum_{j-1}^{i}P_j - \sum_{j=1}^{i}e_j)^2 \tag{18}$$

where:

- r is the number of outcomes which is three in this use case
- P_j is the modelled probability of outcome j
- e_j is the actual probability of outcome j

5.3.3 Configuring the Multi-layer Perceptron Model

As well as the choice of loss metric and use of observed or oversampled data, there many hyperparameters to be tuned in order to configure the optimal model. This study focused on the following parameters:

- Optimisation Function of hidden layers
- Activation Function
- Batch Size
- Initialisation Model
- Weight Constraints
- Dropout Rate
- Number of Epochs
- Model Layer Nodes

A function was written to accept one or many values for each of the parameters listed. In initial tests a wide range of values for the parameters were selected which were then narrowed as the performance was evaluated.

5.3.4 Selected Model

A series of dense fully connected perceptron models were generated and the performance of each was evaluated.

The input layer takes the 8 features described at the beginning of this section. The first hidden layer has 32 neurons before dropout and batch normalisation steps. Two additional hidden layers with 128 and 256 neurons respectively, follow. These layers are also followed by dropout and batch normalisation steps. The output has three neurons to capture the probability of home win, draw and away win results.

The "relu" activation function was used in the hidden layers and the "softmax" activation function was used in the output layer. The model was trained using the Rank Probability Score loss function with a batch size of 100 for 17 epochs. In the selected model the dropout percentage was set to 0, essentially removing the dropout step from the architecture. The input weights and bias weights on each layer of the model were constrained to a maximum value of 2.0.

6 Evaluation, Results and Discussion

6.1 Introduction to Evaluation Methodology

The performance of each model was evaluated by comparing the match outcome probabilities generated to the Betfair market odds. A simple betting strategy was simulated where bets were placed where market odds exceeded the model implied fair value odds by a threshold percentage. The formula below describes the bet decision mathematically:

$$if\ marketOdds > 1 + ((1/modelProb - 1) * (1 + thresholdPercentage))\ then\ Bet \tag{19}$$

The choice of a 25% threshold is arbitrary. The level was chosen as in initial tests 25% maximised profitability. To illustrate the strategy, market odds and model generated probabilities for a match have been presented below.
marketOdds for the home team, Chelsea, to win are 1.58. The right-hand side of the evaluation formula is:

$$= 1 + ((1/0.716 - 1) * (1 + thresholdPercentage)) \tag{20}$$

$$= 1 + ((1/0.716 - 1) * (1.25)) = 1.496 \tag{21}$$

Date	Home Team	Away Team	Home Odds	Draw Odds	Away Odds	Home Prob	Draw Prob	Away Prob
2024-05-05	Chelsea	West Ham	1.58	5.2	5.7	71.6%	15.4%	13.0%

Figure 10: Market Odds and Probabilities used for Evaluation

Therefore odds of 1.496 are deemed the value point where a bet is placed at the 25% threshold level. As odds of 1.58 were on offer prior to the match, the simulated places a bet on Chelsea to win. On this occasion Chelsea won the match and the strategy recorded a 0.58 profit. Had Chelsea lost or the match drawn a 1.0 unit loss would have been recorded.

6.2 Data and Time Periods in Evaluation

The traditional machine learning and neural network models were trained on English Premier League 2015-16 to 2022-23 seasons with the 2023-24 season used for model evaluation. The models selected were those that performed well on the 2023-24 season. The methodology is intended to be broad and applicable to any football league with a liquid market. Therefore, in order to get a true measure of the strength of the approach, data from completely outside of the training and selection dataset was used.

Spain's La Liga and Germany's Bundesliga are comparable to the English Premier League in terms of success, strength and market dynamics. Form features were generated for these two leagues using the methodology and parameter values described in section 4. The models which were selected as optimal when trained and tested on the Premier League were run on these leagues and the results evaluated.

As the Poisson model uses a frequency distribution approach, there are no training and test datasets. Results from all periods have equal importance in performance evaluation. Season average performance results are presented on figure 11.

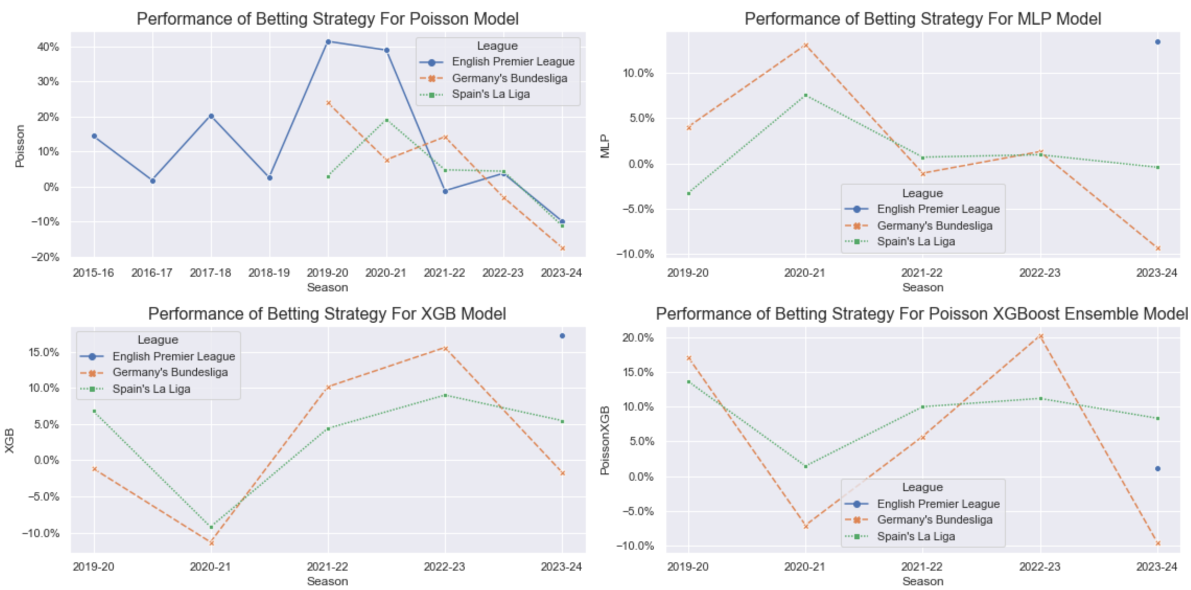

Figure 11: Model Performance Results

6.3 Poisson Distribution Results

The betting strategy based on simulated €1 bets placed where the 25% greater than fair value threshold is met is performed on the three league datasets. On nine seasons of the English Premier League the strategy returns an average of 11.5% yield. The Bundesliga and La Liga have lower average profitability of 5.2% and 4.1% respectively.

6.4 Extreme Gradient Boosting Model Results

The same betting strategy was run on the Extreme Gradient Boosting Model predictions. As the model was trained on earlier seasons, only the 2023-24 English Premier League season is consider for evaluation.

The XGBoost model returns and an impressive 17.2% yield on the English Premier League test season. While still profitable, the model generates less impressive results on the Bundesliga and La Liga with yields of 2.1% and 3.3% respectively.

6.5 Neural Network Results

The same betting strategy was run on the Multi-layer Perceptron model predictions. As the model was trained on earlier seasons, only the 2023-24 English Premier League season is consider for evaluation. Similar to XGBoost, results the neural network model have an impressive 13.5% yield on the English Premier League test season but lower average results on the Bundesliga and La Liga with yields of 1.6% and 1.0% respectively

6.6 Discussion

All four of the models presented have positive average performance results - the simulated betting strategy generates a return. There is a lot of variation within those averages, however.

The Poisson Distribution generates a 41.5% return on the 2019-20 season of the Premier League but returns negative 17.5% on the 2023-24 season of the Bundesliga. The average return is 8.0% over all periods evaluated. Patterns are difficult to discern in the results but recent seasons have shown poorer results than earlier seasons.

The Extreme Gradient Boosting models generated a 17.2% return on the Premier League test season of 2023-24 but less impressive results on the leagues the model was not trained on - 2.1% on the Bundesliga and 3.3% on La Liga for an overall average of 4.2%. Again, patterns are difficult to discern but season results have something approaching an inverse profile to the Poisson results. The Poisson model had large positive returns for the 2020-21 seasons but large negative returns for the 2023-24 seasons. The Extreme Gradient Boosting model large negative returns for the 2020-21 seasons but large positive returns for the 2023-24 seasons.

While generating a 13.5% return on the Premier League test season of 2023-24, the Multi-layer Perceptron model had less impressive results on the leagues the model was not trained on. Returns of 1.6% on the Bundesliga and 1.0% on La Liga for the five seasons reviewed.

An ensemble model was created by combining the probabilities of the Poisson and Extreme Gradient Boosting models. This approach attempts to use signals from both models to generate more robust predictions. Applying the mean probabilities of the models generates profitability in 9 of the 11 seasons reviewed for an average return of 6.5% overall.

7 Conclusion and Future Work

The research question asked if a novel approach to estimating the strength of football teams combined with artificial intelligence techniques could reveal inefficiencies in football betting markets. The profitability shown for the three modelling approaches using the simulated strategies demonstrates that the question has been answered positively. The strategies of each approach generated more than 10% profitability on the English Premier League test data. The models were less profitable on the other two leagues tested. This is to be expected as the form parameters were tuned on, and the machine learning models were trained on historical English Premier League data. Therefore, the positive performance of the simulated strategies provides confirmation of the strength of the methodology.

Testing on Bundesliga and La Liga implicitly assumes that these leagues have the same dynamics. For example, in the preparation of Form statistics the Squad Value Pass Through parameter was set to 0.75. The study showed that estimating team strength at the start of each season using 75% of the squad value performance line and 25% from the previous season performance generated best fitting statistics for the English Premier League. The dynamics may differ for other leagues and lower or higher parameter values may produce better results.

Similarly, the machine learning models trained on English Premier League data may not have optimal configuration or weights for other leagues. Retraining the models on historical data from these leagues would likely produce better performing models.

Acknowledgement

I want to thank Dr. Catherine Mulwa for her guidance and help preparing this research report.

References

Angelini, G., De Angelis, L. and Singleton, C. (2022). Informational efficiency and behaviour within in-play prediction markets, *International Journal of Forecasting* **38**(1): 282–299.

Baboota, R. and Kaur, H. (2019). Predictive analysis and modelling football results using machine learning approach for english premier league, *International Journal of Forecasting* **35**(2): 741–755.

Bondt, W. F. D. and Thaler, R. H. (1989). Anomalies: A mean-reverting walk down wall street, *Journal of Economic Perspectives* **3**(1): 189–202.

Canterbery, E. R. (1999). Irrational exuberance and rational speculative bubbles, *The International Trade Journal* **13**(1): 1–33.

Fama, E. F. (1965). The behavior of stock-market prices, *The journal of Business* **38**(1): 34–105.

Gerhards, J. and Mutz, M. (2017). Who wins the championship? market value and team composition as predictors of success in the top european football leagues, *European Societies* **19**(3): 223–242.

Głowania, S., Kozak, J. and Juszczuk, P. (2022). New voting schemas for heterogeneous ensemble of classifiers in the problem of football results prediction, *Procedia Computer Science* **207**: 3393–3402.

Işın, A. (2023). Home advantage and referee bias: a comparative analysis of league level differences, *Perceptual and Motor Skills* **130**(6): 2621–2631.

Kahneman, D. and Tversky, A. (2013). Prospect theory: An analysis of decision under risk, *Handbook of the fundamentals of financial decision making: Part I*, World Scientific, pp. 99–127.

Kozak, J. and Głowania, S. (2021). Heterogeneous ensembles of classifiers in predicting bundesliga football results, *Procedia Computer Science* **192**: 1573–1582.

Krieger, K. and Davis, J. L. (2022). How much of 'home field advantage'comes from the fans? a natural experiment from the covid-19 pandemic, *UNLV Gaming Research & Review Journal* **26**(1): 6.

Nivetha, S., Geetha, M., Suganthe, R., Prabakaran, R. M., Madhuvanan, S. and Sameer, A. M. (2022). A deep learning framework for football match prediction, *2022 International Conference on Computer Communication and Informatics (ICCCI)*, IEEE, pp. 1–7.

Tiwari, E., Sardar, P. and Jain, S. (2020). Football match result prediction using neural networks and deep learning, *2020 8th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)*, IEEE, pp. 229–231.