National College of Ireland

Configuration Manual

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

National College of Ireland

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:	Configuration Manual
Word Count:	XXX
Page Count:	15

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:	14th July 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):	

1 Introduction

The configuration manual details the system environment and technical specification used to create the research report.

2 Environment Set-Up

Data processing for the project was performed on a machine with the following specifications:

- **Processor:** Intel(R) Core(TM) i5-3570 CPU @ 3.40GHz.
- **Operating System:** Windows 10 Home. 64-bit.
- **Memory:** 16.0 GB
- **Storage:** 446 GB Solid State Hard-drive

3 Processing Environment

Data processing was performed mainly in a Jupyter Notebook version 6.4.8 running Python 3.9.12.

3.1 Python Libraries Used

3.1.1 Data Extraction

The following packages were used to extract data and convert into formats which were more easily readable:

- pickle
- bz2
- json
- os

3.1.2 Data Manipulation

The following packages were used to manipulation and process data:

- pandas
- numpy
- datetime
- copy
- csv
- collections
- imblearn

3.1.3 Statistics and Machine Learning

The following packages were used to calculate statistics and run machine learning algorithms. The Keras and Tensorflow packages are version 2.12.0.

- scipy
- sklearn
- tensorflow
- keras
- sklearn

3.1.4 Data Presentation

The following packages were used to plot charts and present data:

- seaborn
- matplotlib
- tabulate

4 Datasets

4.1 Betfair Market Data

Betfair Exchange provides historical betting transaction data on various markets it covers including football markets. The data is made available to subscribers on its historical data portal [1]. The data is provided in a zipped binary file for each football match analysed in the study. Python scripts searched these files and extracted betting transaction data which was then saved to flat files.

4.2 Football Result and Statistics

Football result and statistics data was downloaded from the Football.co.uk [2] website in CSV format. Each league season is presented on one file. 25 season files were download for 11 seasons of the English Premier League, 7 seasons of the German Bundesliga and 7 seasons of Spain's La Liga.

5 Process and Code Description

Data processing in the project was performed mainly in the Jupyter Notebook environment running Python code. Microsoft Excel was used in data validation, to calculate profitability and to presentation results. The sections below describe the Jupyter Notebooks used and present justifications for decisions made and parameter value selections during the project.

5.1 1_CalcHomeAdvantage

The script loads 11 seasons of English Premier League results. It calculates simple and exponentially weighted rolling averages of the home team advantage over the seasons. It

[1] https://historicdata.betfair.com/

[2] https://www.football-data.co.uk/data.php

Table 1: Description of formula variables used throughout the project

Variable Name	**Description of Form Variable**
HTGSForm	Home Team Goal Score Form
HTGCForm	Home Team Goal Concede Form
ATGSForm	Away Team Goal Score Form
ATGCForm	Away Team Goal Concede Form
HomeAdv	Home Advantage Goals
projHomeGoals	Projected Home Goals Scored in match
projAwayGoals	Projected Away Goals Scored in match
deltaHomeGoals	Observed minus projected Home Goals Scored in match
deltaAwayGoals	Observed minus projected Away Goals Scored in match
xGHome	Expected Goals statistic for home team in match
xGAway	Expected Goals statistic for away team in match
xGWeight	Weight assigned to xG performance metric
homeGoals	Actual Goals Scored for home team in match
awayGoals	Actual Goals Scored for away team in match
formAdjProp	Proportion of team delta to adjust form statistics by

compares the averages to the observed home advantage of the next match over the period covered and calculates the errors for each average. Results are presented is chart and table format.

Figure 1 presents a rolling 380 match simple rolling average of the home advantage metric. 380 matches were chosen as the averaging window as it includes a whole season.

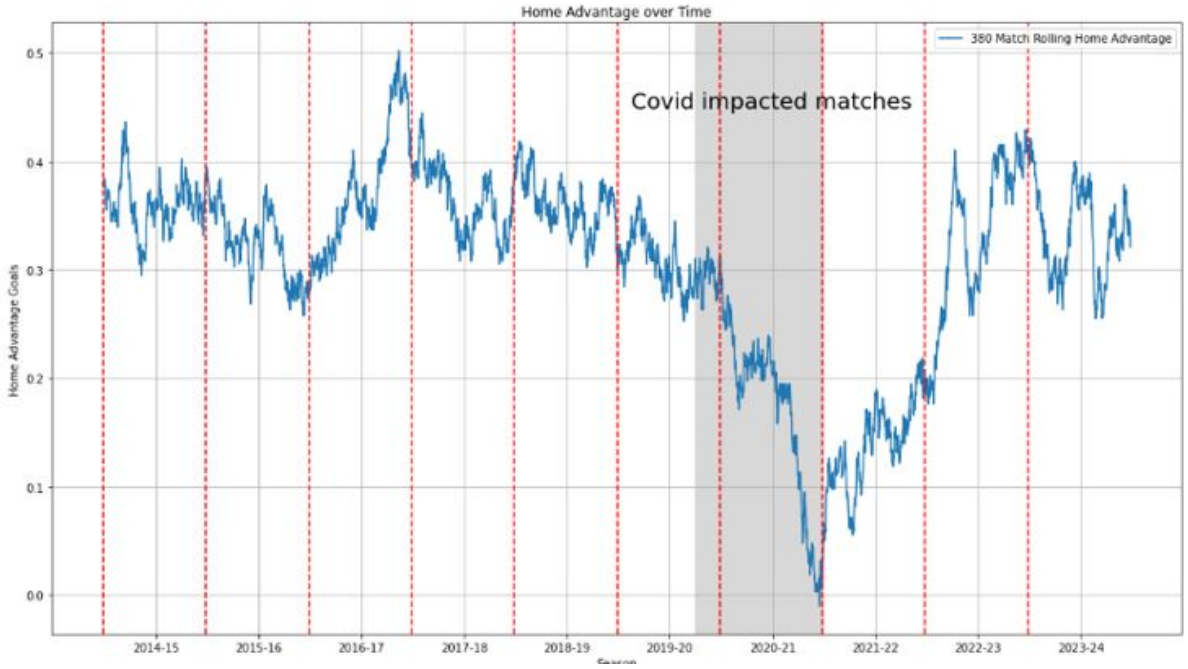

Figure 1: Rolling Home Advantage from 2014-15 to 2023-24

The vertical red lines indicate season breaks. The shaded area covers matches at the end of the 2019-20 and the entire 2020-21 season which were impacted by the Covid 19

pandemic. In March 2020 the English Premier League was suspended [3]. When matches resumed in June of that year they took place in empty stadiums bar a small number team management and staff. In September the 2020-21 season began with empty stadiums once again. The entire season was played with no or very few spectators allowed in the stadiums. The chart shows that the lack of spectators or some other reason led to a dramatic fall in the home advantage.

Table 2 presents the results of a comparison of different averaging methods and observer home advantage in scorelines.Smaller errors will indicate an average number which better tracks actual goal difference.

Table 2: Calculated errors of different averaging types

Averaging Method	**Root Mean Squared**	**Mean Absolute**
380 match simple average	1.897518	1.477106
190 match simple average	1.905402	1.485181
95 match simple average	1.911899	1.488717
47 match simple average	1.920689	1.494916
380 match exponentially weighted	1.901070	1.483786
190 match exponentially weighted	1.904032	1.484902
95 match exponentially weighted	1.909609	1.487766
47 match exponentially weighted	1.920576	1.494911

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.

5.2 2_CalcSeasonAveragesStats

The notebook loads the historical match results and statistics. It calculates the average performance of each team for the seasons covered. It measures performance as combination of actual goals scored/conceded and the expected goals for/expected goals against performance metric. Different weightings of actual and expected goals are tested. The differences between observed results and averaged performance at the various weightings are measured in order to identify the weigth which minimises errors.

5.2.1 Averaging approach to measure team performance over season

The average performance for each team over the season was measured using the formulas:

$$avgGoalsFor = \sum_{1}^{n}(GoalsScored_n * actWeight + xGFor_n * (1 - actWeight)) \quad (1)$$

$$avgGoalsAgain = \sum_{1}^{n}(GoalsConcede_n * actWeight + xGAgain_n * (1 - actWeight)) \quad (2)$$

where:

[3] https://www.premierleague.com/news/1682374

- n is the number of matches played
- GoalsScored and GoalsConcede are the number of goals scored and goals conceded by the team being evaluated in a particular match
- xGFor and xGAgain are performance metrics for the particular match. These metrics measure how many goals the team would be expected to score (xGFor) and concede (xGAgain) given the performance level and number of scoring opportunities created.
- actWeight is the proportion of the performance measure attributed to the actual number of goals scored. The remaining proportion (1 – actWeight) is the proportion of the performance measure attributed to the xG statistics.

GoalsScored, GoalsConceded, xGFor and xGAgain are observers or captured statistics. actWeight is a weight parameter which was tuned to find the optimal value. Figure 2 presents a line plot for the error results.

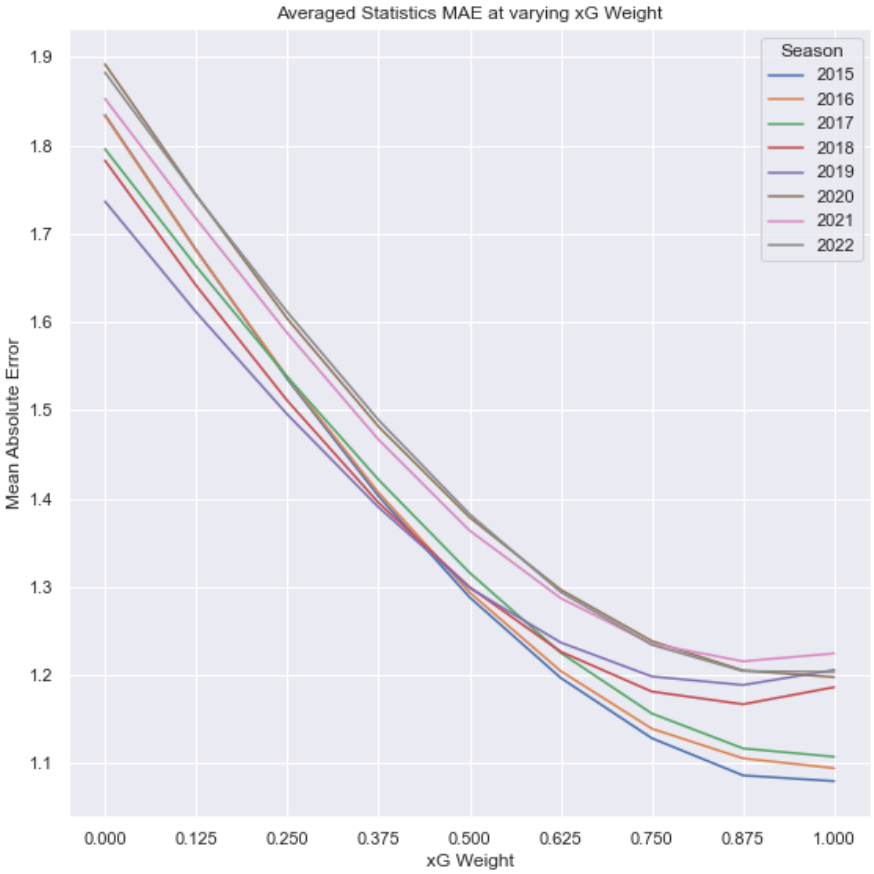

Figure 2: Mean Average Errors of averaged form by xGWeight parameter values

5.3 3_CalcOptimisedForm and 3_Oth_CalcOptimisedForm

These notebooks load the historical match results and statistics and uses an iterative approach to calculate the optimal goal score and goal concede form statistics in each season. The first notebook is for the English Premier League and the second for Spain's La

Liga and Germany's Bundesliga. In both notebooks average form numbers are calculated for each team as the fist step. Then the goal form statistic is adjusted up slightly for a team and the delta between projected scorelines and observed scorelines in measured. If the delta increases the adjustment is discarded and the form is adjusted down slightly and the delta is again measured. This process is repeated for each team in the league over three cycles. The process iteratively team strength or form measures that better reflect team performance over the season.

The iterative process was run for the 2015-16 to 2023-24 seasons for a range of xG-Weight and formAdjustProp parameters. Figure 3 plots the Mean Average Errors identified for the parameters tested across the seasons.

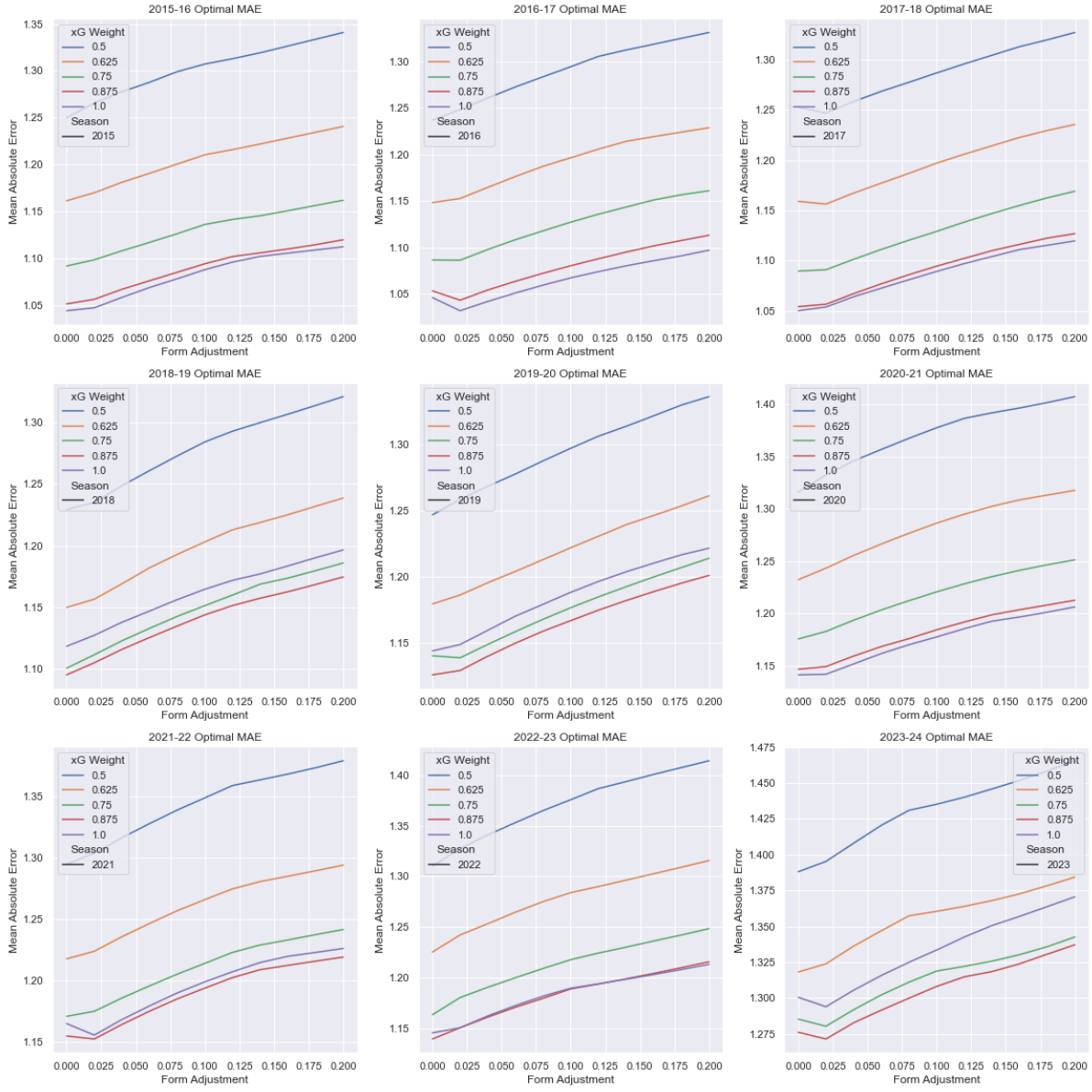

Figure 3: Mean Average Errors of optimised form using a range of xGWeight and formAdjustProp parameter values

5.4 4_FormAdjustmentsOffS and 4_Oth_FormAdjustmentsOffS

These notebooks load the optimised season form generated in the previous notebooks and team squad values for each team. The first notebook is for the English Premier League and the second for Spain's La Liga and Germany's Bundesliga. Ordinary least squares regression lines are calculated for each season to demonstrate that team performance is related to the value of the team squad. Start season form numbers are calculated for each team in each season by combining the previous season form and the regression line expected team form given the value of the squad. A range of start season start Form numbers are generated by applying different previous season and squad value regression line weightings.

5.4.1 Relationship between squad value and team "Form" metrics

Figure 4 shows goals scored and goals conceded by each team plotted against squad value. Goals stats are from the 2019-20 season and squad values are at the season end. Ordinary least squares regression lines have been added to the charts. This regression line can be considered as the expected performance level given the quality of the players in the squad. During the season Bournemouth and Manchester City scored more goals than would be expected by the value of their squads while Huddersfield, Tottenham and Chelsea scored fewer. In the Goals Conceded chart teams below the regression line have conceded fewer goals and therefore outperformed their squad value. During the 2019-20 season, Wolverhampton Wanderers considerably outperformed their squad value while Fulham and Manchester United under-performed their squad value in terms of goals conceded.

The explained variance statistics for each season are presented in Table 3. The average of these figures is 63.6% showing that approaching two thirds of the performance of teams in the English Premier League can be explained by the value of the team's squads.

Table 3: Explained Variance squad value on team performance

Explained Variance of Goals Scored & Conceded against Squad Value									
Season	2014-2015	2015-2016	2016-v17	2017-2018	2018-2019	2019-2020	2020-2021	2021-2022	2022-2023
Scored	0.7470	0.4172	0.8299	0.7335	0.8174	0.7152	0.6931	0.7773	0.4155
Conceded	0.5843	0.6326	0.7012	0.7241	0.6430	0.4649	0.6048	0.5252	0.4166

5.5 5_BetfairDataProcessing

This notebook opens a series of .bz2 files on a folder, searches for relevant football match betting transactions which are then saved to flat files. The data is cleaned and missing data match data added where necessary. The results are csv files saved for each league season. The files contain the price of the last home team, draw and away team bet transacted prior to the match start time.

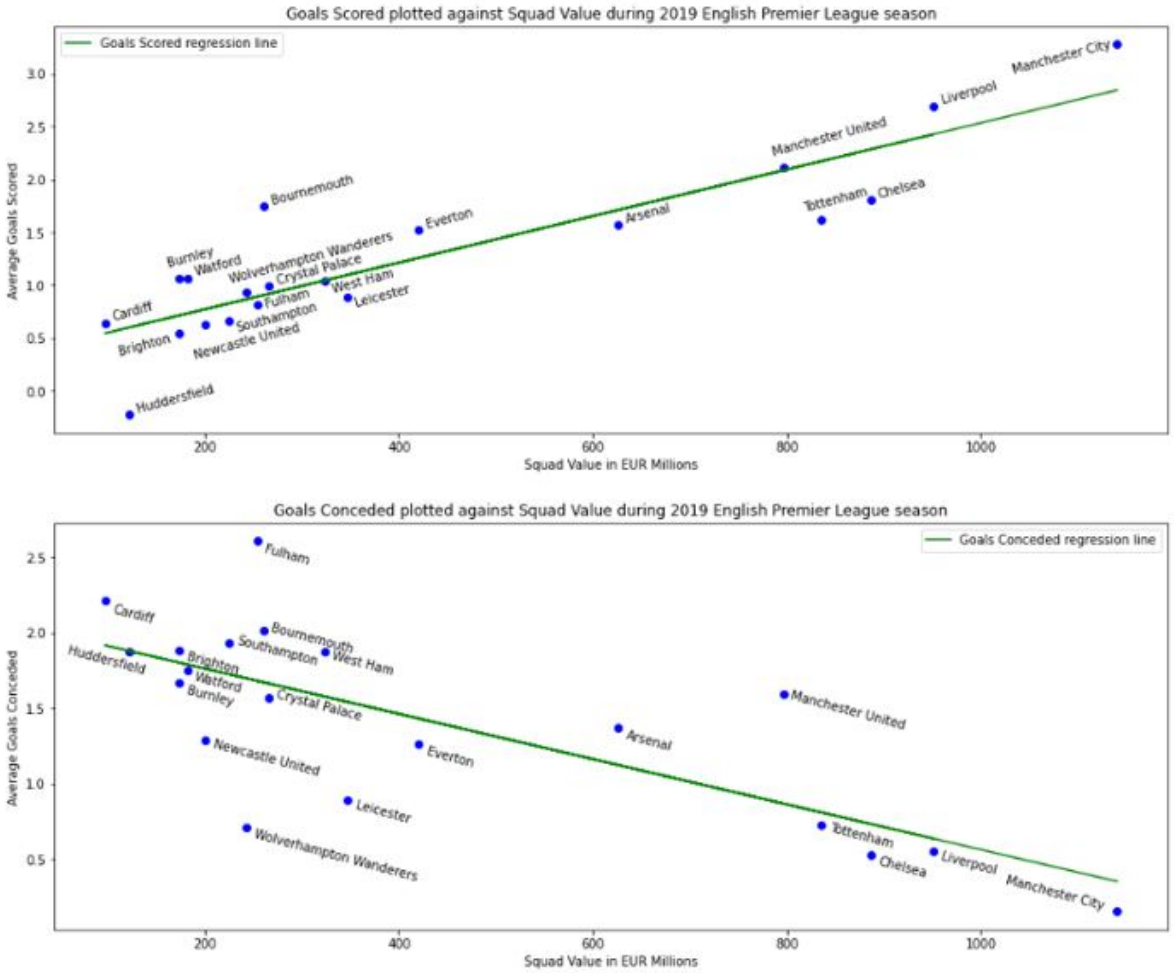

Figure 4: Relationship between squad value and team performance

5.6 6_SeasonFormEvolution and 6_Oth_SeasonFormEvolution

These notebooks load the historical match results and statistics for each season along with the start of season form values calculated for each team in the previous notebooks. For each set of start season form values and using a range of xGWeightProp, formAdjProp and squadValuePassThrough variable values, an evolution of the season is run. A delta value is calculated for each set of parameters in each season in order to identify which set of parameter most closely predict the actual scorelines.

Figure 5 presents boxplots of delta values against 0.75, 0.875 and 1.0 xG Weight values. Delta values calculated for lower xG Weights are higher than those presented. The delta values are average of the mean absolute error between projected and actual scoreline over a season. Delta figures presented in figure 9 are averages of the numbers calculates over seasons 2015-16 to 2022-23. The plots on the left have been generated using the averaged form approach described in section 5.2.1. The plots on the right show values generated using the form optimisation approach.

Figure 6 shows that for both the averaging and optimisation approaches, errors are minimised at the 0.875 xG Weight proportion. The study will therefore focus on estimates of form using this parameter. Figure 10 below presents delta values for the averaging and optimisation approaches at the 0.875 xG Weight level. The x-axis is the Form Adjustment proportion and each line presents a different Squad Pass Through level.

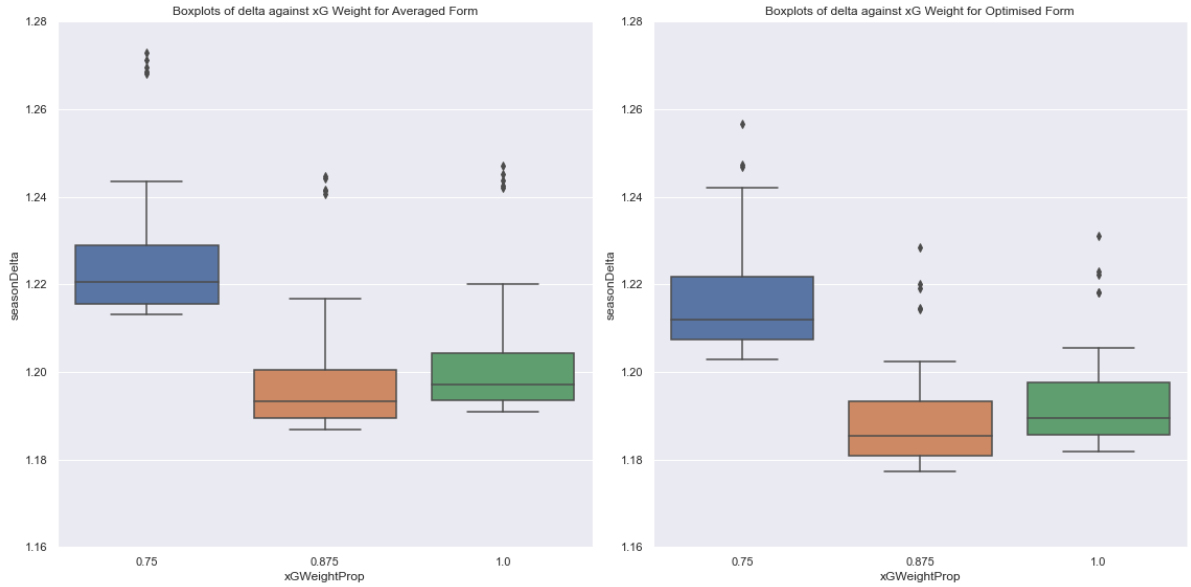

Figure 5: Boxplots of delta values against the best performing xG Weight values

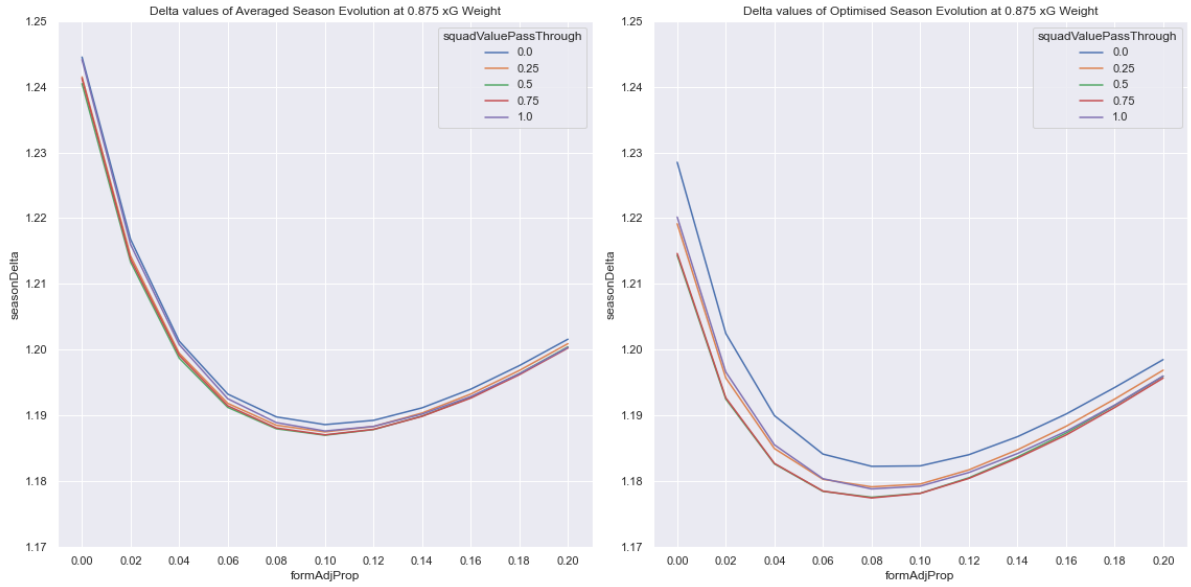

Figure 6: Plots of delta values against Form Adjustment for different Squad Value Pass Through values

5.7 7A_PredictionModel_Poisson

This notebook loads the historical form statistics with the best fitting parameters as determined in earlier parts of the study. Using the entire match history, the study calculates how closely a Poisson distribution curve aligns to historical results. An adjustment ratio is calculated for each point in the distribution. For example, using a Poisson curve to predict the number of goals scored by the away team with λ set to the mean number of goals will understate the frequency of 0 goals scored and overstate the frequency of 1 goal scored. The adjustment ratios correct for these deviations.

Figure 7 illustrates the Poisson distribution curve and adjustments required in order to create a frequency distribution which more closely reflects observed distributions.

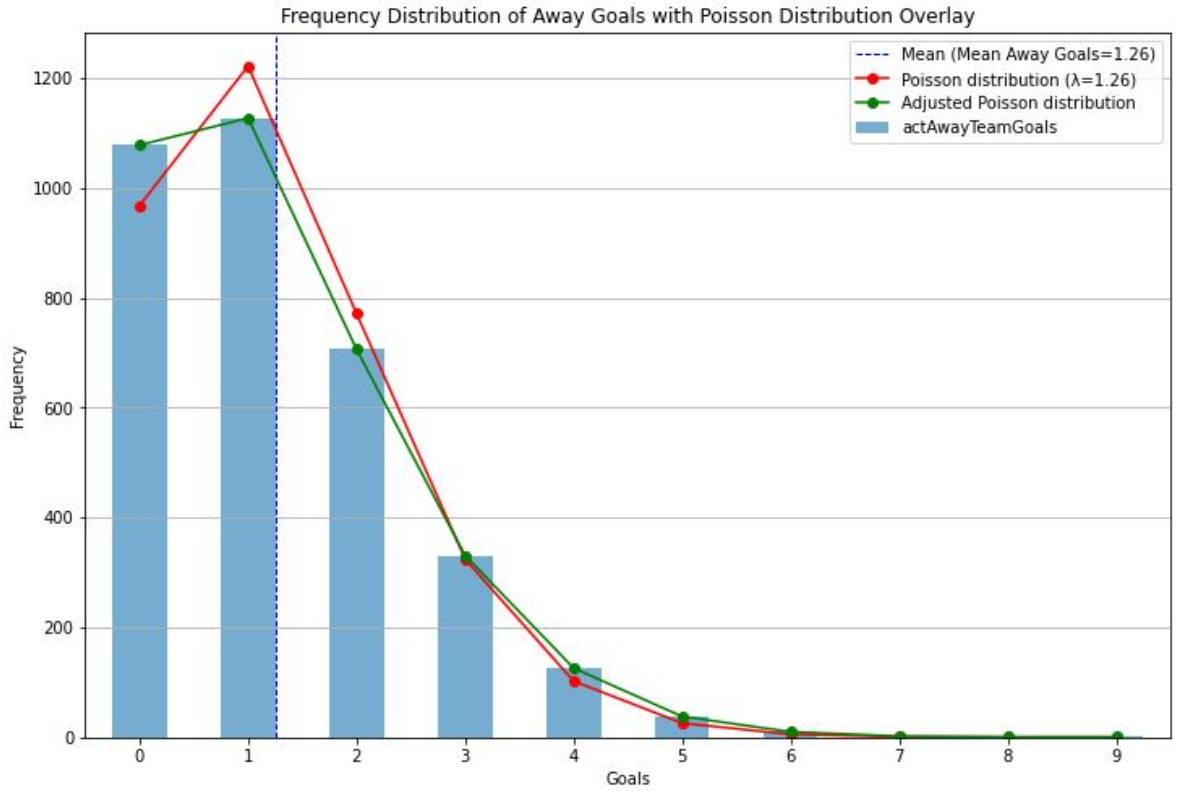

Figure 7: Away Goals Frequency with Poisson and Adjusted Poisson Distributions

This notebook includes a function to calculate the cross probabilities and sum probabilities of scoreline for each outcome. The individual goal score probabilities, i.e. the frequency or probability of each team scoring 0 to 6 goals is calculated for each match in the dataset. The cross probability function is then employed to convert the individual score probabilities into result probabilities.

5.8 7B_PredictionModel_Traditional

This notebook loads the historical form statistics and market odds data. Using a grid-search type approach, it searchs a wide array of parameters in each of algorithm types. 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 models.

The model performance was evaluated by generating probabilities for each match in the held out 2023-24 season. Match result probabilities generated by the models were saved to a CSV file to be reviewed against the market odds in order to determined the profitability or otherwise of each model.

5.9 7C_PredictionModel_MLP

This notebook again loads the historical form statistics and market odds data. In preparation for neural network algorithm a Synthetic Minority Oversampling Technique or SMOTE algorithm was applied to the test dataset in order to balance the classes. The SMOTE technique generates new records of the minority classes – the away win and draw

results – using a nearest neighbours' approach. Figure 8 presents the number of records before and after SMOTE was applied to the dataset.

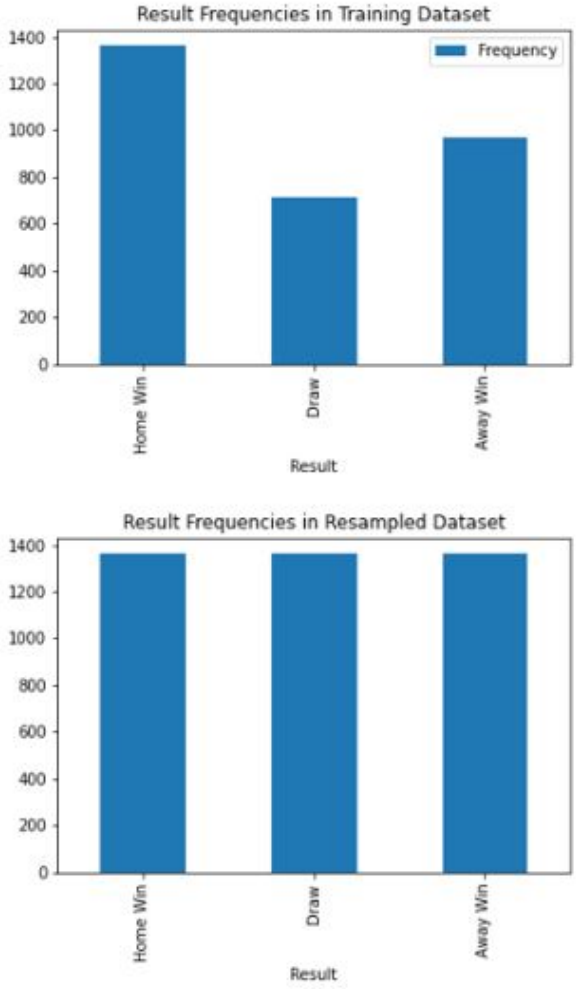

Figure 8: Number of records before and after SMOTE was applied to the training dataset

The best forming model was identified as having the configuration presented in figure 9.

5.10 8_EnsembleModel and 8_Oth_EnsembleModel

This 8_EnsembleModel notebook loads the historical form statistics and market odds data for the English Premier League. The 8_Oth_EnsembleModel notebook performers the same processes and calculations on Spain's La Liga and Germany's Bundesliga datasets.

The Poisson Distribution model predictions are run again in these notebooks. The best performing XG Boost and Multi-layer Perceptron models and weights are loaded also. A set of model probabilities are then generated for each of the three algorithms. The result probabilities implied by the market odds are calculated and considered a fourth set of result probabilities.

The four sets of results are combined in all possible permutations. These are then analysed in the Excel LeagueModelReview.xlsm file to determine if a combination of models can outperform the individual models.

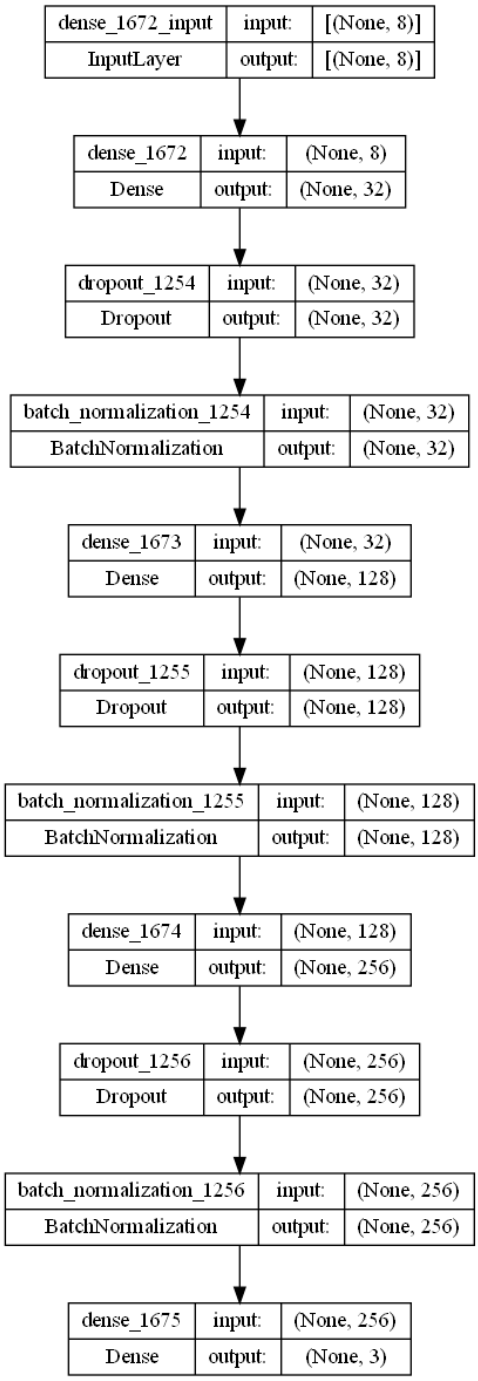

Figure 9: Architecture of selected Multi-layer Perceptron model

6 Results

6.1 LeagueModelReview.xlsm

Model output probabilities as well as the Betfair market odds are loaded to the spreadsheet for review and analysis. In each case the analysis takes the same approach. Model prediction probabilities are compared to the market odds of bets transacted prior to the matches and a simulated betting strategy is employed. Where market odds are 25% higher than the fair value odds implied by the model probabilities, a €1 bet is placed.

The spreadsheet calculates the number of bets place, the profit or loss in money and percentage terms. The numbers are summarised by league and season.

6.2 Poisson Distribution Results

Figure 10 below presents the results of the Poisson distribution model evaluated on the three leagues. The P&L values are based on simulated €1 bets placed where the 25% greater than fair value threshold is met. All three leagues show this model is profitable on average. However, the model generated large losses for the 2023-24 season for each league.

English Premier League				P&L			Count of bets					Profitability
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2015-16	-€11	-€2	€44	€31	81	29	102	212	-14.2%	-5.5%	42.8%	14.4%
2016-17	€21	€28	-€43	€5	126	61	107	294	16.7%	45.2%	-40.4%	1.8%
2017-18	€32	€12	€2	€46	106	40	83	229	30.3%	31.1%	2.2%	20.3%
2018-19	€45	-€15	-€24	€6	106	29	79	214	42.1%	-53.1%	-30.1%	2.6%
2019-20	€29	€10	€65	€104	104	37	109	250	27.5%	28.1%	59.4%	41.5%
2020-21	€7	€14	€84	€104	108	37	123	268	6.4%	37.6%	68.0%	39.0%
2021-22	-€7	€3	€1	-€4	55	38	198	291	-13.3%	7.9%	0.4%	-1.2%
2022-23	€10	€14	-€16	€9	118	18	96	232	8.8%	78.9%	-16.5%	3.8%
2023-24	-€24	-€6	-€3	-€33	126	51	156	333	-19.3%	-11.0%	-2.0%	-9.9%
Total	€101	€59	€109	€268	930	340	1053	2323	10.8%	17.3%	10.3%	11.5%

Germany's Bundesliga				P&L			Count of bets					Profitability
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	-€2	€16	€37	€51	85	32	96	213	-2.9%	50.3%	39.0%	24.0%
2020-21	€24	€1	-€9	€16	49	39	122	210	49.1%	1.5%	-7.1%	7.6%
2021-22	€37	-€8	€4	€33	136	35	59	230	27.3%	-24.0%	6.7%	14.2%
2022-23	€5	€9	-€19	-€5	81	26	67	174	6.3%	33.1%	-28.5%	-3.1%
2023-24	-€10	-€11	-€19	-€40	145	27	55	227	-6.6%	-41.5%	-34.5%	-17.5%
Total	€54	€6	-€5	€55	496	159	399	1054	10.9%	3.6%	-1.3%	5.2%

Spain's La Liga				P&L			Count of bets					Profitability
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	€22	-€5	-€7	€9	153	37	125	315	14.1%	-13.0%	-5.9%	3.0%
2020-21	-€4	€10	€44	€50	123	42	97	262	-2.9%	24.0%	45.0%	19.1%
2021-22	-€2	€31	-€17	€12	73	22	155	250	-2.3%	140.0%	-11.1%	4.8%
2022-23	€22	-€5	-€6	€10	85	20	135	240	25.4%	-26.0%	-4.4%	4.4%
2023-24	-€29	€17	-€15	-€27	105	17	122	244	-27.6%	97.1%	-12.3%	-11.3%
Total	€9	€47	-€2	€54	539	138	634	1311	1.7%	34.3%	-0.3%	4.1%

Figure 10: Table of Poisson model performance

6.3 Extreme Gradient Boosting Model Results

Figure 11 presents the results of the Extreme Gradient Boosting model evaluated on the three leagues. The P&L values are based on simulated €1 bets placed where the 25% greater than fair value threshold is met. Only the 2023-24 English Premier League season is presented as earlier season were used to train the model. While still profitable, the model generates less impressive results on the Bundesliga and La Liga.

6.4 Neural Network Results

Figure 12 below presents the results of the Multi-layer Perceptron model evaluated on the three leagues. The P&L values are based on simulated €1 bets placed where the

English Premier League	P&L				Count of bets				Profitability			
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2023-24	€26	€44	-€1	€69	158	112	130	400	16.4%	39.0%	-0.6%	17.2%
Total	€26	€44	-€1	€69	158	112	130	400	16.4%	39.0%	-0.6%	17.2%

Germany's Bundesliga	P&L				Count of bets				Profitability			
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	-€20	-€1	€18	-€4	144	74	92	310	-14.2%	-1.1%	19.3%	-1.1%
2020-21	€18	-€17	-€39	-€38	130	87	115	332	13.7%	-19.3%	-33.5%	-11.3%
2021-22	€8	€11	€13	€31	156	56	95	307	5.0%	18.9%	13.6%	10.2%
2022-23	€27	€7	€14	€49	133	79	101	313	20.4%	9.2%	14.3%	15.6%
2023-24	-€5	€8	-€8	-€5	160	57	111	328	-3.1%	13.4%	-7.3%	-1.7%
Total	€27	€8	-€2	€34	723	353	514	1590	3.8%	2.2%	-0.3%	2.1%

Spain's La Liga	P&L				Count of bets				Profitability			
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	€1	€22	€5	€28	148	141	118	407	0.4%	15.7%	4.3%	6.8%
2020-21	-€13	-€13	-€12	-€38	135	108	172	415	-9.9%	-12.3%	-6.7%	-9.2%
2021-22	€4	€4	€9	€18	135	109	157	401	3.3%	3.8%	5.8%	4.4%
2022-23	€24	-€6	€19	€37	147	108	157	412	16.3%	-5.4%	12.1%	9.0%
2023-24	€11	€14	-€3	€22	193	103	102	398	5.7%	13.8%	-3.3%	5.5%
Total	€27	€21	€18	€66	758	569	706	2033	3.5%	3.8%	2.6%	3.3%

Figure 11: Table of XGBoost model performance

25% greater than fair value threshold is met. Only the 2023-24 English Premier League season is presented as earlier season were used to train the model. While still profitable, the model generates less impressive results on the Bundesliga and La Liga.

English Premier League	P&L				Count of bets				Profitability			
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2023-24	€4	€20	€24	€49	141	59	163	363	3.1%	34.2%	15.0%	13.5%
Total	€4	€20	€24	€49	141	59	163	363	3.1%	34.2%	15.0%	13.5%

Germany's Bundesliga	P&L				Count of bets				Profitability			
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	-€16	-€15	€43	€13	133	42	137	312	-12.1%	-35.0%	31.7%	4.1%
2020-21	€38	€10	-€9	€38	88	57	148	293	42.6%	17.7%	-6.2%	13.1%
2021-22	€40	-€16	-€27	-€3	171	38	74	283	23.2%	-42.6%	-35.9%	-1.1%
2022-23	€26	-€19	-€3	€4	125	51	97	273	20.6%	-37.5%	-3.1%	1.3%
2023-24	-€4	-€9	-€15	-€29	191	32	83	306	-2.1%	-28.6%	-18.6%	-9.3%
Total	€83	-€49	-€11	€23	708	220	539	1467	11.7%	-22.3%	-2.0%	1.6%

Spain's La Liga	P&L				Count of bets				Profitability			
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	-€26	€9	€5	-€11	147	101	99	347	-17.4%	8.6%	5.5%	-3.3%
2020-21	€7	€25	-€7	€25	100	108	120	328	6.8%	22.7%	-5.5%	7.5%
2021-22	-€3	€15	-€10	€2	133	94	136	363	-1.9%	16.0%	-7.3%	0.7%
2022-23	€15	-€12	€0	€3	116	106	121	343	12.7%	-11.2%	0.3%	1.0%
2023-24	-€2	€12	-€11	-€1	159	78	107	344	-1.3%	15.0%	-10.5%	-0.4%
Total	-€9	€48	-€22	€18	655	487	583	1725	-1.3%	9.9%	-3.7%	1.0%

Figure 12: Table of Multi-layer Perceptron model performance

6.5 Ensemble Model Results

Figure 13 presents the results of an ensemble model. Probabilities used are the mean of the Poisson and Extreme Gradient Boosting model probabilities. This approach attempts to use signals from both models to generate more robust predictions.

English Premier League				P&L			Count of bets				Profitability		
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2023-24	-€10	€36	-€22	€4	131	82	106	319	-7.8%	43.7%	-20.4%	1.2%
Total	-€10	€36	-€22	€4	131	82	106	319	-7.8%	43.7%	-20.4%	1.2%

Germany's Bundesliga				P&L			Count of bets				Profitability		
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	-€14	€10	€45	€41	109	47	83	239	-13.1%	22.0%	54.0%	17.1%
2020-21	€23	€5	-€47	-€19	82	65	117	264	27.9%	8.0%	-39.9%	-7.0%
2021-22	€19	€7	-€12	€14	133	38	74	245	14.6%	17.1%	-16.0%	5.7%
2022-23	€25	€5	€20	€50	110	56	83	249	22.8%	9.2%	24.2%	20.2%
2023-24	-€10	€4	-€20	-€26	157	41	78	276	-6.6%	9.7%	-25.3%	-9.5%
Total	€43	€31	-€13	€61	591	247	435	1273	7.2%	12.6%	-3.1%	4.8%

Spain's La Liga				P&L			Count of bets				Profitability		
Season	Home	Draw	Away	Total	Home	Draw	Away	Total	Home	Draw	Away	Total
2019-20	€22	€10	€13	€45	139	98	94	331	16.0%	10.5%	13.5%	13.6%
2020-21	-€2	-€3	€9	€5	120	87	135	342	-1.4%	-3.2%	7.0%	1.5%
2021-22	-€11	€28	€18	€35	107	83	155	345	-10.3%	33.7%	11.3%	10.0%
2022-23	€24	-€2	€13	€35	104	70	140	314	22.9%	-3.1%	9.6%	11.2%
2023-24	€9	€12	€7	€28	164	74	98	336	5.2%	16.4%	7.6%	8.4%
Total	€42	€45	€60	€148	634	412	622	1668	6.6%	11.0%	9.7%	8.9%

Figure 13: Table of Poisson and Extreme Gradient Boosting ensemble model performance