MENDEL
Soft Computing Journal

Predicting Football Match Outcomes With Machine Learning Approaches

Bing Shen Choi, Lee Kien Foo✉, Sook-Ling Chua

Multimedia University - MMU Cyberjaya, Malaysia

lkfoo@mmu.edu.my✉

Abstract

The increasing use of data-driven approaches has led to the development of models to predict football match outcomes. However, predicting match outcomes accurately remains a challenge due to the sport's inherent unpredictability. In this study, we have investigated the usage of different machine learning models in predicting the outcome of English Premier League matches. We assessed the performance of random forest, logistic regression, linear support vector classifier and extreme gradient boosting models for binary and multiclass classification. These models are trained with datasets obtained using different sampling techniques. The result showed that the models performed better when trained with dataset obtained using a balanced sampling technique for binary classification. Additionally, the models' predictions were evaluated by conducting simulation on football betting profits based on the 2022-2023 EPL season. The model achieved the highest accuracy is the binary class random forest, but the model provided the highest football betting profit is the binary class logistic regression.

1 Introduction

Like many other competitive sports, football has become very data-driven. This trend has led to an increasing interest in research on football match outcome prediction. The ability to predict football matches outcomes holds significant importance to football clubs, as it allows them to extract valuable insights and gain a competitive edge over their opponents. [4] demonstrated the significant potential to increase the understanding of football team performance by embracing a data-driven approach. Similarly, individuals involved in football betting are leveraging the abundance of football data to predict football matches in the hopes of gaining profit.

One of the biggest challenges in predicting football matches outcomes is the imbalanced nature of the dataset, as there are relatively more football matches that end in win for the home team. A study by [14] revealed that the home advantage phenomenon exists in many team sports, including football. In order to address this imbalanced issue, we have examined the impact of training data selected using different sampling techniques together with different set-up of target classes i.e. binary vs. multiclass in this research.

The scope of this research only extends to predicting the outcomes of football matches within the English Premier League (EPL), which is one of the most renowned and competitive football leagues globally. Throughout a season in the EPL, each team will play every other team in the league twice (home and away). Matches can result in a home team win, a draw or a home team loss. As there are 20 teams in the EPL, each team will play 38 games (19 home games 19 away games) in a season. At the end of the season, the team with the most points wins the league. In this paper, the terms "Win", "Draw", and "Lose" are used in relation to the home team. Specifically, a "Win" refers to a victory for the home team or a loss for the away team, while a "Lose" refers to a loss for the home team or a win for the away team. This research also evaluates the profits generated from football betting during the 2022-2023 EPL season based on our best model.

2 Literature Review

Prediction tasks in sports using machine learning are often treated as a classification problem [3]. Football is no different as most researchers try to predict each match as win, draw or lose, which is a multiclass classification problem. However, one challenge faced by many researchers who performed multiclass classification is the difficulty in predicting draws. [15] demonstrated that victories and defeats are highly reflected in the team's performance during a match, but draws are challenging to predict using a machine learning approach. [6] noted that their long short-term memory model was not predicting any of the draws, while [12] reported the same finding for their logistic regression model. [2] trained 4 models, including Gaussian naive bayes, support vector machine, random forest, and XGBoost, and reported that each model underperformed in predicting draws. According to [6], this could be caused by imbalanced data, as is usually the case with

a relatively lower number of draws in the dataset.

Alternatively, the problem can be framed as a binary classification. [16] trained their model on home wins and away wins, achieving an accuracy of 69.5% on data spanning five seasons. [17] compared multiclass and binary classification. In the case of binary classification, they classified the matches into home win and non-home win. They discovered that using logistic regression for binary classification resulted in a slightly higher accuracy (77.43%) compared to multiclass classification (70.27%).

The studies in this area also emphasise the importance of feature engineering. [2] highlighted that only 2 out of 12 features they used were not engineered features, suggesting that feature engineering plays a critical role in predicting football matches. Football match statistics data are commonly used for feature engineering. [1] and [18] computed the aggregate value of match statistic indicators for both the home and the away teams based on previous games using either average or sum. Researchers also commonly use FIFA ratings to quantify the footballing ability of the teams [2, 6, 16, 18, 19, 20, 21]. FIFA ratings are numerical values developed by EA Sports that represent the skills and abilities of players and teams. The company hires scouts all over the world to rate players and teams as realistically as possible each year [5]. [16] used only FIFA ratings for the home and away teams as features and achieved an impressive accuracy of 69.5% using logistic regression, highlighting the reliability of FIFA data in predicting match outcomes. Other interesting features such as month is used by [7] as they thought the weather might have an impact on the match, while [22] used the number of days from a team's last match as an indicator of the players' fatigue level.

A common focus in this field of research is the prediction of football with the aim of beating the bookmakers. In addition to using the common classification evaluation metrics, [18] evaluated their models based on the amount of profit gained from football betting. Their most successful model was the random forest technique, achieving an accuracy of 65.26% and a profit margin of 26.74%. In contrast, [2] fell short of outperforming the bookmakers' predictions, although they did achieve promising results with an accuracy of around 60%.

In this study, we compared the performance of models trained on datasets obtained using different sampling techniques for both binary and multiclass classification. We then assessed our models in the context of football betting by simulating the betting profits using matches from the most recent season of the EPL.

3 Research Methodology

3.1 Dataset

The data used in this research comprises the EPL matches spanning 10 seasons, starting from 2012-2013 till 2021-2022. The data was collected from various

Table 1: Target variable distribution in the dataset.

Match Outcome of Home Team	Number of Records
Win	1694
Draw	903
Lose	1203
Total	3800

online data sources. The football-data.co.uk [10] and fbref.com [8] provide historical football matches statistics, while fifaindex.com [9] provides FIFA ratings of players and teams that prove useful in assessing the ability of the football teams. We also manually collected data for the home stadium (venues) and the historical league rankings of each team. After performing data pre-processing, the datasets from these various data sources were integrated into one dataset, where each row represents one single historical football match. Table 1 shows the number of records for each match outcome in the dataset.

3.2 Feature Engineering

In football matches prediction tasks, it is important to only use features that could be known before the start of each match [11], so that the model can be used to make prediction before the football match take place. From the collected data, we have combined the relevant features for both the home team and away team by taking the difference between the values. For instance, we combined the Overall Rating of the home team and the Overall Rating of the away team by taking the difference between them. In other words, the gap of the Overall Rating between the home team and the away team is the feature that will be used in our model building. Table 2 shows all the features that were obtained after feature engineering.

In addition to the 26 features shown in Table 2, the 4 features shown in Table 3 were also included in our dataset.

Since a team's performance in the previous matches is a good indicator of how well they would perform in the next match. We have engineered our features by taking the average value of 5 matches prior to the current match. This means that the first 5 matches of the season for any given team will not be used as data points. Table 4 shows the number of records in the final dataset after feature engineering.

3.3 Feature Selection

Inspired by the work done by [18], we have researched on two methods to perform feature selection. In Feature Selection Method 1, we first computed the correlation matrix between all the features. This allows us to identify features that are highly correlated with each other and select only one of them. This step eliminated 14 features, leaving us with 16 features left.
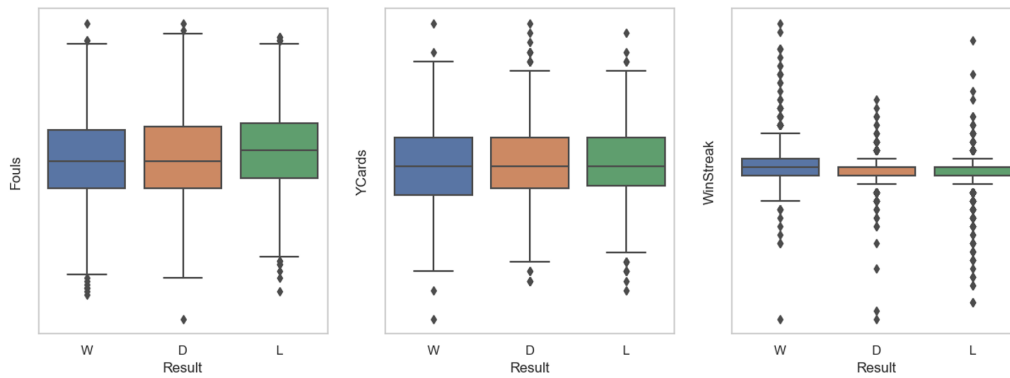
Figure 1: Examples of variables with significant overlapping boxplots.

Table 2: Features that were created from feature engineering.

| No.	Features	Data Type	Description
1	Round	Numerical	Difference between the round of the game being played during the season.
2	Days	Numerical	Difference between the number of days since the last match played
3	Shots	Numerical	Difference between the average number of shots taken during the previous 5 matches
4	ShotsOT	Numerical	Difference between the average number of shots on target taken during the previous 5 matches
5	Corners	Numerical	Difference between the average number of corners taken during the previous 5 matches
6	Fouls	Numerical	Difference between the average number of fouls committed during the previous 5 matches
7	YCards	Numerical	Difference between the average number of yellow cards received during the previous 5 matches
8	RCards	Numerical	Difference between the average number of red cards received during the previous 5 matches
9	GoalsScored	Numerical	Difference between the average number of goals scored during the previous 5 matches
10	GoalsConceded	Numerical	Difference between the average number of goals conceded during the previous 5 matches
11	GoalsDiff	Numerical	Difference between the average goals difference during the previous 5 matches
12	HTGoalsScored	Numerical	Difference between the average number of goals scored by half-time during the previous 5 matches
13	HTGoalsConceded	Numerical	Difference between the average number of goals conceded by half-time during the previous 5 matches
14	HTGoalsDiff	Numerical	Difference between the average goals difference by half-time during the previous 5 matches
15	Points	Numerical	Difference between the average points gained during the previous 5 matches
16	WinPercent	Numerical	Difference between the win percentage during the previous 5 matches
17	WinStreak	Numerical	Difference between the win streak in terms of number of matches
18	UnbPercent	Numerical	Difference between the unbeaten percentage during the previous 5 matches
19	UnbStreak	Numerical	Difference between the unbeaten streak in terms of number of matches
20	Def	Numerical	Difference between the team's defending ability represented in the form of FIFA rating
21	Mid	Numerical	Difference between the team's midfield ability represented in the form of FIFA rating
22	Att	Numerical	Difference between the team's attack ability represented in the form of FIFA rating
23	Ovr	Numerical	Difference between the team's overall ability represented in the form of FIFA rating
24	LastSeasonRank	Numerical	Difference between the team's last season league ranking
25	WinOdd	Numerical	Difference between the B365 betting odd for the Home Team and Away Team win
26	PromotedMatchup	Categorical	The type of matchups: 0 = Both teams are not recently promoted team or are promoted teams 1 = Home Team is a promoted team -1 = Away Team is a promoted team
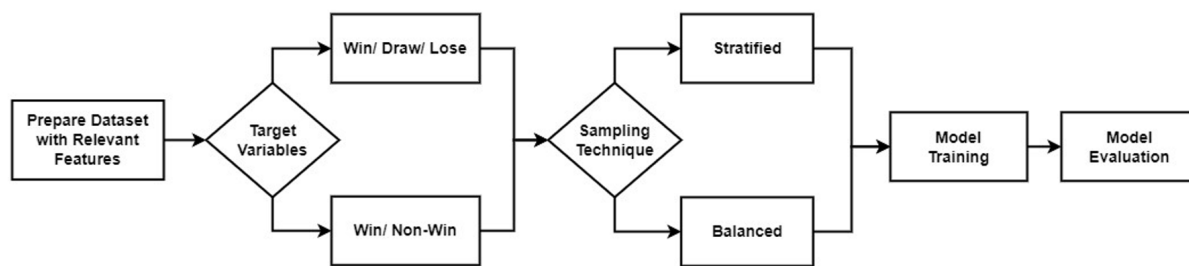
Figure 2: An overview of the experimental process.

Next, we plotted a boxplot between each remaining feature with the match outcome. The objective is to remove features where the boxes have significant overlap as shown in Fig. 1. A significant overlap between the boxes suggests that the feature does not contribute much in identifying the match outcome. This step eliminated 6 features, leaving us with 10 features. Lastly, we used the BORUTA feature selection method [13] on the remaining features to identify those that have a significant impact on the target variable. BORUTA uses the random forest algorithm to compare the relevance of original features with their corresponding "shadow features". In the end, the BORUTA algorithm eliminated 2 features, leaving us with 8 features (Feature Set 1). Table 5 displays all the features that are in Feature Set 1.

The difference between Feature Selection Method 1 and Feature Selection Method 2 is the order of the process. In Feature Selection Method 2, we first used the BORUTA algorithm on all the features, then removed the highly correlated features in the feature set determined by the BORUTA algorithm, and lastly eliminating features with significant overlapping boxes. Feature Selection Method 2 eventually left us with only 6 features which are Shots, Corners, Points, UnbStreak, LastSeasonRank and Venue (Feature Set 2).

All the 6 features in Feature Set 2 are already included in the Feature Set 1. Feature Set 1 has 2 additional features which are Referee and PromotedMatchup. Hence, we decided to proceed with Feature Set 1 in our experiments.

4 Experimental Setup

In this study, we have performed classification on multiclass prediction and binary class prediction. For multiclass prediction, the target is to predict the match outcome on "Win", "Draw" or "Loss". For binary class prediction, we have grouped the "Draw" and "Loss" classes into a single class label as "Non-Win".

4.1 Sampling Techniques

For each of the binary class and multiclass prediction, we have selected the training data by using two different sampling techniques, the stratified and the balanced sampling techniques. For stratified sampling, the training data is the combination of the 80% data

selected from each class. The number of records of each class in the training data is thus proportionate to the number of records of each class in the original dataset.

For balanced sampling, we have selected 80% data from the class with the smallest size and then matched the records from the other classes. For example, in the multiclass prediction, the class with the smallest data size is "Loss", so we selected 618 (80%) records from the "Loss" class. The same number of records (618) will be selected from the "Win" and the "Draw" class. All the selected records are then combined to form the training data while the remaining records are used for testing. The training data selected using the balanced sampling technique thus has lesser number of records as compared to training data selected using the stratified sampling technique.

We then trained different classification models on each of the resulted training data. A 5-fold cross validation is carried out in each experiment. Fig. 2 shows the overview of the experiments that were conducted.

4.2 Classification Models

We have trained 4 classification models: random forest, logistic regression, linear support vector classifier and extreme gradient boosting. All models are trained for binary and multiclass classification.

4.2.1 Random Forest

Random forest (RF) classifier is an ensemble learning algorithm that combines the predictions of multiple decision trees to make predictions. Each decision tree is built independently on a subset of the training data, which is created through a technique called bootstrap sampling. When the RF classifier makes predictions, each decision tree would independently predict the class label of a data instance and the final prediction would be the class that is predicted the most number of times.

4.2.2 Logistic Regression

Logistic regression (LR) models the relationship between the input features and the probability of a data instance belonging to a particular class. It uses a sigmoid function to transform the linear combination of input features and weights into a value between 0 and

Table 3: Features that were not created from feature engineering.

No.	Features	Data Type	Description
1	Referee	Categorical	The match referee
2	Day Of Week	Categorical	The day of the week that the match takes place (Sunday, Monday etc.)
3	DrawOdd	Numerical	B365 betting odd for a draw
4	Venue	Categorical	The stadium in which the match is played

Table 4: Target variable distribution after feature engineering.

Match Outcome of Home Team	Number of Records
Win	1496
Draw	772
Lose	1029
Total	3297

1, which is then mapped to the probability of the data instance belonging to the positive class. Hence, to classify a data instance, a decision boundary is needed. Instances with a predicted probability above the decision boundary are classified as the positive class, while those below are classified as the negative class. In a multiclass classification problem, the LR uses the "one-vs-rest" strategy. For each class, a binary LR classifier is trained to distinguish that class from the rest. This results in multiple LR classifiers, each representing a specific class against all other classes. To make predictions on a data instance, the class with the highest predicted probability is assigned as the predicted class.

4.2.3 Linear Support Vector Classifier

Linear support vector classifier (SVC) is a variant of support vector machines. Linear SVC finds an optimal hyperplane that separates the data instances of binary classes in the feature space. The optimal hyperplane is the one that maximizes the distance between the hyperplane and the closest data points of each class. It applies a linear kernel function to perform classification. Similar to LR, the linear SVC uses the "one-vs-rest" strategy to solve a multiclass classification problem.

4.2.4 Extreme Gradient Boosting

Like RF, the extreme gradient boosting (XGBoost) is a decision tree ensemble technique. XGBoost works on the principle of boosting, which involves training multiple decision trees sequentially, where each subsequent decision tree focuses on correcting the errors made by the previous decision trees. The predictions are then made by aggregating the predictions of all the decision trees in the ensemble.

4.2.5 Model Evaluation

The performance of all the models in our experiment is evaluated with accuracy, F1 score and area under the ROC curve (AUC). The accuracy is a metric that measures the proportion of correctly classified instances out of the total number of instances in a dataset.

$$accuracy = \frac{(TP + TN)}{(TP + FP + TN + FN)} \quad (1)$$

The F1 score is a metric that combines the precision and recall of a model. Precision quantifies the ratio of true positive predictions to all positive predictions made by the model, while recall quantifies the ratio of true positive predictions to all actual positive instances in the dataset.

$$precision = \frac{TP}{TP + FP} \quad (2)$$

$$recall = \frac{TP}{TP + FN} \quad (3)$$

$$F1Score = 2 \times \frac{precision \times recall}{precision + recall} \quad (4)$$

The AUC is commonly used to evaluate the performance of binary classification models. It measures the trade-off between the true positive rate and the false positive rate at different classification thresholds by plotting the ROC curve. The AUC score is calculated by the area under the ROC curve. A high AUC score indicates that the model is good at distinguishing between the positive and negative classes. For multiclass classification, the "one-vs-rest" strategy is used.

4.3 Football Betting Simulation

From the results of the experiments conducted using multiclass classification (Win / Draw / Lose) and binary classification (Win / Non-Win), we selected the best performing models to simulate football betting profits by predicting the football matches of the 2022-2023 EPL season. The betting odds used is provided by Bet365.

For the simulation, we started with 100 units of money. For simulation using multiclass models, 1 unit of money is invested in each match following the model's prediction. If the model's prediction is correct, the profit will be calculated by the formula: (betting odds - 1). However, if the model's prediction is incorrect, we lose the unit money invested and the profit will be -1, which means a loss of one unit of money. In the case of simulating betting profits using binary class models, 1 unit of money will be invested only when the model predicts a Win. If the model predicts a Non-Win, no unit of money will be invested into either the draw odds or lose odds.

Table 5: The features in Feature Set 1.

| Feature	Description
Referee	The match referee
Shots	The difference between the average number of shots taken during the previous 5 matches by home team and away team
Corners	The difference between the average number of corners taken during the previous 5 matches by home team and away team
Points	The difference between the average number of points gained during the previous 5 matches by home team and away team
UnbStreak	The difference between the number of unbeaten match streak of home team and away team
LastSeasonRank	The difference between the last season league ranking of home team and away team
Venue	The stadium in which the match is played
PromotedMatchup	The type of matchups: 0 = Both teams are not recently promoted team or are promoted teams 1 = Home Team is a promoted team while away team is not -1= Away Team is a promoted team while home team is not
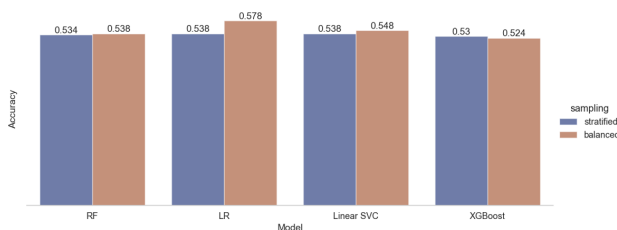
Figure 3: Accuracy of Models Trained on Win/Draw/Lose
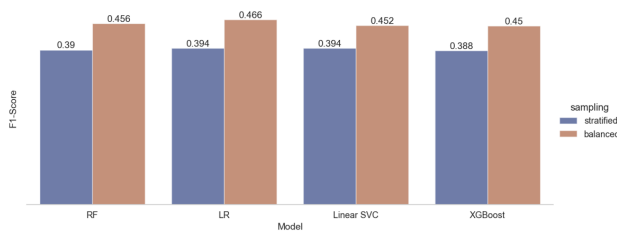

Figure 4: F1-Score of Models Trained on Win/Draw/Lose

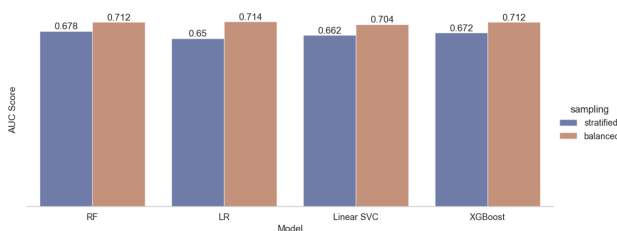Figure 5: AUC Score of Models Trained on Win/Draw/Lose

5 Results and Discussions

5.1 Multiclass Model

Fig. 3, 4 and 5 show the comparison of accuracy, F1 score and AUC score for all four models trained on multiclass classification with training data selected using stratified sampling or balanced sampling.

Across all four models, the models trained on data selected with balanced sampling performed better in all 3 metrics, except for accuracy in XGBoost. When comparing among the models, LR performed the best with the highest accuracy, F1 score and AUC score. Further investigation in the result showed that all four models that were trained on training data selected using stratified sampling technique failed to predict "Draw", while the models trained on training data selected using balanced sampling technique were able to predict a few "Draws".

5.2 Binary Class Model

Fig. 6, 7 and 8 show the comparison of the performance for all four models trained on binary classification with training data selected using stratified sampling or balanced sampling.

There is a mixed result in the accuracy and F1 score for models trained on data selected using the two sampling techniques, although all the models performed better in AUC score when trained with data selected using the balanced sampling technique. When comparing among models, RF trained with data selected using the balanced sampling technique performed the best. When comparing multiclass models to binary class models, all the binary class models performed better. Thus, it can be concluded that by combining "Draw" and "Loss" into one class, we can improve the classification performance of the models.

5.3 Football Betting Simulation

Based on the results of the previous experiments, LR trained on data selected with balanced sampling technique perform the best for multiclass prediction, while RF trained on data selected with balanced sampling technique has best performance for binary class prediction. Thus, we chose these 2 models for the football betting simulation. In this experiment, both models are re-trained with hyperparameter tuning for multiclass and binary class prediction. Table 6 shows the

details of these models.

After training the models, we used the model to predict the 2022-2023 EPL season. We started with 100 units of money. Table 7 shows the betting profits achieved by each model. Positive return is achieved when we simulate the betting with the binary class LR or RF model while the multiclass models generated a loss in our simulation.
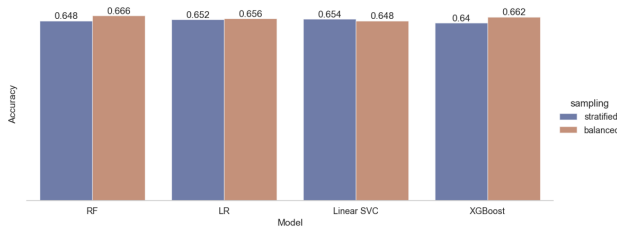

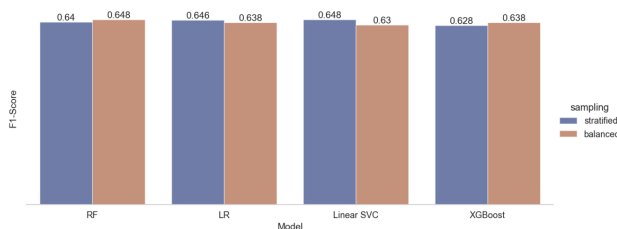Figure 6: Accuracy of Models Trained on Win/Non-Win
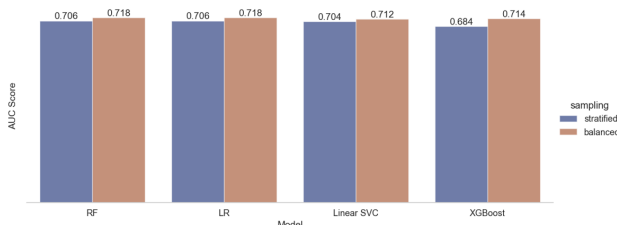

Figure 7: F1-Score of Models Trained on Win/Non-Win

Figure 8: AUC Score of Models Trained on Win/Non-Win

5.4 Further Discussion

The conducted experiments revealed several key findings. Firstly, the models that were trained on data selected using balanced sampling technique performed better than the models that were trained on data selected using stratified sampling technique. All the multiclass models that were trained on data selected using stratified sampling technique could not predict the "Draw" class at all. This may be due to the inherent imbalance nature of the dataset, which causes the models to fail to learn the "Draw" class. We have conducted a secondary experiment that compares the performance of models trained on a dataset of "Draw" and "Non-Draw" by combining the "Win" and "Loss" into a single class. All the models performed poorly with accuracy less than 50%. This shows that the models

are incapable of differentiating draws in this particular dataset. Thus, when the "Draw" and "Loss" classes are combined into a single class, the models performed better as the dataset becomes less imbalanced and the models become less biased. Our simulation on football betting also showed that multiclass models are not reliable for football betting as they generate negative returns. The models that were trained on binary classification ("Win" and "Non-Win") are more practical in football betting.

5.5 Limitation

Most of the features that were used were derived by taking the average of the matches statistics in the previous 5 games. As a result, a limitation arose wherein the constructed models are unable to make predictions for the first 5 matches of the season, since the models require features that are derived from the previous 5 matches. In any given EPL season, the models would only be able to perform prediction on the 6th match onwards.

Like many research in this area, we quantified the ability of the team using ratings from FIFA video game series, but it would be more accurate to only account for the ability of the starting players in a particular match. However, getting data on the starting players in each historical football match is challenging and time consuming. Besides, instead of using ratings from the FIFA video game series, the idea of quantifying team and player abilities using ratings from the Football Manager video game series can also be considered.

6 Conclusion

In this research, we have experimented on applying machine learning algorithms in predicting football matches outcomes. A key finding is that models that were trained with data selected using balanced sampling technique performed better than models that were trained with data selected using stratified sampling technique, as the former were able to predict draws. The experiments also revealed that the multiclass logistic regression model has the best performance among all the multiclass models, while the binary class random forest model has the best performance among all binary models. Lastly, we have simulated football betting profits using our logistic regression and random forest models. The highest return was achieved by the binary class logistic regression model, which successfully generated 7.57% returns. For future enhancement, we would like to explore using historical match data from the lower tier English football leagues. By including the lower tier English football leagues, the data size will increase and may lead to an improvement in the model performance. We have performed classification with four well known machine learning methods in this research, we are keen to explore deep learning methods for football matches outcome prediction in next stage.

References

[1] ANFILETS, S., BEZOBRAZOV, S., GOLOVKO, V., SACHENKO, A., KOMAR, M., DOLNY, R., KASYANIK, V., BYKOVYY, P., MIKHNO, E., AND OSOLINSKYI, O. Deep multilayer neural network for predicting the winner of football matches. *International Journal of Computing 19*, 1 (2020), 70–77.

[2] BABOOTA, R., AND KAUR, H. Predictive analysis and modelling football results using machine learning approach for english premier league. *International Journal of Forecasting 35*, 2 (2019), 741–755.

[3] BUNKER, R. P., AND THABTAH, F. A machine learning framework for sport result prediction. *Applied computing and informatics 15*, 1 (2019), 27–33.

[4] CINTIA, P., GIANNOTTI, F., PAPPALARDO, L., PEDRESCHI, D., AND MALVALDI, M. The harsh rule of the goals: Data-driven performance indicators for football teams. In *2015 IEEE International Conference on Data Science and Advanced Analytics (DSAA)* (2015), IEEE, pp. 1–10.

[5] COTTA, L., DE MELO, P., BENEVENUTO, F., AND LOUREIRO, A. Using fifa soccer video game data for soccer analytics. In *Workshop on large scale sports analytics* (2016).

[6] DANISIK, N., LACKO, P., AND FARKAS, M. Football match prediction using players attributes. In *2018 World Symposium on Digital Intelligence for Systems and Machines (DISA)* (2018), IEEE, pp. 201–206.

[7] ELMILIGI, H., AND SAAD, S. Predicting the outcome of soccer matches using machine learning and statistical analysis. In *2022 IEEE 12th Annual Computing and Communication Workshop and Conference (CCWC)* (2022), IEEE, pp. 1–8.

[8] FBREF.COM. Football statistics and history. https://fbref.com/en/ (accessed 01 October 2023).

[9] FIFA. Fifa index. https://www.fifaindex.com/ (accessed 01 October 2023).

[10] FOOTBALL DATA. Football results, statistics & soccer betting odds data. https://football-data.co.uk/ (accessed 01 October 2023).

[11] GOMES, J., PORTELA, F., AND SANTOS, M. F. Decision support system for predicting football game result. In *Computers-19th International Conference on Circuits, Systems, Communications and Computers-Intelligent Systems and Applications Special Sessions. Series* (2015), vol. 32, pp. 348–353.

[12] IGIRI, C. P., AND NWACHUKWU, E. O. An improved prediction system for football a match result. *IOSR Journal of Engineering* (2014).

[13] KURSA, M. B., AND RUDNICKI, W. R. Feature selection with the boruta package. *Journal of statistical software 36* (2010), 1–13.

[14] LEITE, W. S. Home advantage: Comparison between the major european football leagues. *Athens Journal of Sports 4*, 1 (2017), 65–74.

[15] PAPPALARDO, L., AND CINTIA, P. Quantifying the relation between performance and success in soccer. *Advances in Complex Systems 21*, 03n04 (2018), 1750014.

[16] PRASETIO, D., ET AL. Predicting football match results with logistic regression. In *2016 International Conference On Advanced Informatics: Concepts, Theory And Application (ICAICTA)* (2016), IEEE, pp. 1–5.

[17] RAJU, M. A., MIA, M. S., SAYED, M. A., AND UDDIN, M. R. Predicting the outcome of english premier league matches using machine learning. In *2020 2nd International Conference on Sustainable Technologies for Industry 4.0 (STI)* (2020), IEEE, pp. 1–6.

[18] RODRIGUES, F., AND PINTO, Â. Prediction of football match results with machine learning. *Procedia Computer Science 204* (2022), 463–470.

[19] RUDRAPAL, D., BORO, S., SRIVASTAVA, J., AND SINGH, S. A deep learning approach to predict football match result. In *Computational Intelligence in Data Mining: Proceedings of the International Conference on ICCIDM 2018* (2020), Springer, pp. 93–99.

[20] SAIEDY, S., QACHMAS, M., AND AMANULLAH, F. Predicting epl football matches resutls using machine learning algorithms. *International Journal of Engineering Applied Sciences and Technology 5* (2020), 83–91.

[21] STÜBINGER, J., MANGOLD, B., AND KNOLL, J. Machine learning in football betting: Prediction of match results based on player characteristics. *Applied Sciences 10*, 1 (2019), 46.

[22] TAX, N., AND JOUSTRA, Y. Predicting the dutch football competition using public data: A machine learning approach. *Transactions on knowledge and data engineering 10*, 10 (2015), 1–13.