National College of Ireland

Higher Diploma in Science in Computing (HDAIML_SEP24OL)

2024/25

Adrian Tarin Martinez

x23388978

A Machine Learning Approach for Predicting Football Match Outcomes

Cross Assessment Project

1 INTRODUCTION

Predicting football match outcomes using machine learning has been explored from various angles. Classification models, dynamic team strength ratings, deep learning, and betting market profitability. Football is the world's most popular sport, with a global betting market exceeding €30 billion annually. Predicting match outcomes is a high-impact challenge, as bookmakers rely on proprietary models to set odds, often leading to structural inefficiencies in pricing [1]. This project develops a proof-of-concept (PoC) machine learning pipeline to predict football match outcomes using hybrid data (match statistics and betting odds), aiming to benchmark against bookmaker models and identify arbitrage opportunities.

This proof-of-concept (PoC) establishes a focused framework for football match prediction by integrating three critical dimensions. First, it combines traditional match statistics with historical betting odds, creating a hybrid data approach that captures both team performance metrics and market sentiment signals. Second, the project prioritizes model transparency by employing interpretable machine learning techniques, directly addressing the prevalent "black-box" limitations of conventional bookmaker systems. These interconnected components collectively lay the groundwork for a comprehensive future project capable of standardizing odds evaluation methodologies, refining betting strategies through data-driven insights, and ultimately promoting greater fairness and transparency in sports betting markets.

The sports betting industry currently faces significant challenges stemming from inconsistent odds compilation practices across bookmakers. As demonstrated in [2], these discrepancies arise from varying risk models, diverse data sources, and inherent market biases, creating both arbitrage opportunities and systemic trust issues. A transparent machine learning solution could address these problems by systematically identifying market inefficiencies for traders while reducing the industry's dependence on opaque proprietary systems. Academic research in this domain reveals three persistent gaps: most existing models fail to effectively combine statistical and odds data, rarely validate predictions through simulated betting scenarios, and typically lack interpretability features that would enable practical adoption by stakeholders. For the author, this project represents a strategic synthesis of professional sports betting experience and technical machine learning skills, with direct relevance to career aspirations in sports trading or machine learning engineering roles where data-driven odds compilation is increasingly valued.

The investigation centers on three pivotal questions that structure the research approach: First, which classification algorithm demonstrates superior accuracy in predicting match outcomes among the tested alternatives? Second, to what extent do betting odds enhance model performance when compared to predictions based solely on basic match statistics? Third, and most critically, can the optimal model configuration generate consistent positive returns when benchmarked against established bookmaker odds? These questions collectively address both the technical efficacy and practical utility of the developed prediction system.

2 LITERATURE REVIEW

Predicting football match outcomes using machine learning has been extensively studied, with approaches ranging from classification models to dynamic team ratings and betting market integration. This section critically evaluates key works, highlighting their strengths, limitations, and relevance to this project.

2.1 Predictive Modeling Approaches

Bunker and Thabtah [3] treat football prediction as a classification task, emphasizing the use of Artificial Neural Networks (ANNs) and temporal data handling via round-by- round prediction. Their CRISP-DM-based framework discourages cross-validation in favor of training-test splits that reflect chronological match progression.

Stubinger and Knoll [1] benchmark multiple models on complex football datasets and show that Random Forests (RF) consistently outperform other models like SVM and logistic regression—both in accuracy and in generating actual profits.

Tammouch et al. [4] focus on dimensionality reduction and apply DNN, RF, and Decision Trees to a dataset with over 6,000 records, similar to ours. Their use of PCA and Mutual Information to streamline features parallels our own methodology.

Constantinou and colleagues introduce a family of models based on relative scoring and probabilistic ratings. Their Pi-ratings [5], and later the Dolores model [6], rank teams dynamically and demonstrate consistent profitability over five Premier League seasons. These ratings are often more predictive than traditional metrics and odds.

Luiz et al. [7] show that football is indeed predictable when deep neural networks are fed domain-specific features like relative attack/midfield/defense powers. These outperformed even pi-ratings in some contexts, highlighting the power of custom-engineered inputs.

Tax and Youstra [8] apply structured ML methods to the Dutch Eredivisie, testing classifiers with dimensionality reduction. Their hybrid model (LogitBoost + ReliefF) achieved 56.05% accuracy, showing that public data combined with odds can match or beat standalone bookmaker models. These results validate the viability of open data-driven prediction systems.

Wunderlich and Memmert [9] contribute a Betting Odds ELO model that combines the transparency of ELO ratings with the predictive strength of market odds. Their finding— that odds contain more relevant pre-match information than match results themselves—provides a compelling case for fusing statistical models with bookmaker data. Their model outperformed classic ELO, and opens new doors for both forecasting and performance analysis.

The pi-football model by Constantinou et al. [10] leverages Bayesian networks to incorporate subjective variables like managerial impact and match pressure. Used in real-time for EPL matches, it matched bookmaker accuracy and generated profit under fair-market odds, proving its potential in live betting scenarios.

2.2 Market-Aware Models and Betting Strategy

Hubaček et al. [11] minimize correlation with bookmaker odds to exploit market inefficiencies, and apply CNNs for player-level data aggregation. Their use of Markowitz portfolio theory to optimize bet sizes aligns with betting as a form of investment.

Barbosa da Costa et al. [12] shift to the niche BTTS market, showing that under-explored betting categories can be more predictable and profitable with the right features and league-specific tuning.

Walsh and Joshi [13] argue that model calibration, not just accuracy, is key for betting. Their calibrated models delivered up to +36.93% ROI, proving that sharp probability estimation outpaces mere correctness in real-world profitability.

Strumbelj and Robnik-Šikonja [2] analyze odds as probabilistic forecasts, concluding that not all bookmakers are equal. Their odds forecasting quality varies by league and has improved over time, validating odds as a solid but imperfect baseline.

2.3 Key Insights Across Literature

Several recurring themes emerge across the reviewed body of work. Hybrid models, such as those combining ELO ratings with bookmaker odds or public data with betting information, consistently outperform models relying on a single data source. Bookmaker odds themselves are not only competitive but often contain richer predictive signals than historical match results alone. Moreover, evaluation metrics focused on probabilistic calibration and betting correlation tend to offer more informative assessments than simple accuracy scores. It is also evident that league-specific characteristics significantly influence predictability; certain leagues like the English Premier League (EPL) or the Eredivisie exhibit more consistent patterns than others. In addition, handcrafted features and statistical rating systems—such as Pi-ratings, relative team strengths, and odds-based ELO models—play a critical role in enhancing model interpretability and performance.

Collectively, these findings reflect a broader shift in football analytics: from basic classification tasks to the development of calibrated, market-aware, and explainable prediction frameworks. This study builds upon that tradition by proposing a hybrid methodology that integrates open source data, engineered variables, and market signals to produce robust forecasts with potential utility for informed betting strategies.

3 DATA SOURCING AND HANDLING

3.1 Dataset Description

The dataset used in this project was sourced from football-data.co.uk, which provides structured CSV files covering historical match data and betting odds for the top five European football leagues—namely the English Premier League, Spanish La Liga, Italian Serie A, German Bundesliga, and French Ligue 1. The dataset spans six full seasons, from 2018/2019 to 2023/2024, and contains both match-level statistics and market betting odds from various bookmakers.

To understand the scale of the dataset, it's important to note the league structures: the English Premier League, Spanish La Liga, and Italian Serie A each consist of 20 teams. In these leagues, each team plays 38 rounds, with 10 matches per round, resulting in a total of 380 matches per season per league. Meanwhile, the German Bundesliga and French Ligue 1 each feature 18 teams, playing 34 rounds of 9 matches per round, totaling 306 matches per season per league.

Over six full seasons, this leads to a cumulative total of 10,512 matches across all five leagues—forming the complete dataset, where each row corresponds to a single match. Each match entry includes over 80 features spanning a wide array of categories, such as team performance metrics, match outcomes, and pre-match betting odds. The depth and granularity of this data make it well-suited for large-scale predictive modeling and advanced machine learning techniques. These are grouped into the following key categories:

- Match Metadata:
  - o Match date and time
  - o League division
  - o Home and away team names
- Match Outcome Variables:
  - o Full-time and half-time goals for both teams (FTHG, FTAG, HTHG, HTAG)
  - o Full-time result (FTR) coded as H (Home win), D (Draw), or A (Away win)
- Match Statistics (where available):
  - o Shots (HS, AS), Shots on Target (HST, AST)
  - o Fouls, Corners, Offsides, Yellow/Red Cards
  - o Bookings points and other in-game events
- Betting Odds:
  - o Pre-closing odds from over a dozen bookmakers, including Bet365, Pinnacle, and Betfair
  - o Odds for various markets: 1X2 outcomes, Over/Under 2.5 goals, and Asian Handicap
  - o Aggregated statistics such as market average and maximum odds (e.g., AvgH, MaxA)
- Target Variable:
  - The primary response variable used for prediction is the Full-Time Result (FTR), treated as a multi-class classification problem (Home Win, Draw, Away Win).

The richness of the dataset allows for a multi- dimensional analysis of team performance, game dynamics, and betting market behavior. Additionally, the availability of odds from different bookmakers enables the evaluation of market efficiency and the development of model-based betting strategies.

	A	B	C	D	E	F	G	H	I	J
1	Div	Date	Time	HomeTeai	AwayTear	FTHG	FTAG	FTR	HTHG	HTAG
2	SP1	#########	18:30	Almeria	Vallecano	0	2	A	0	2
3	SP1	#########	21:00	Sevilla	Valencia	1	2	A	0	0
4	SP1	#########	16:00	Sociedad	Girona	1	1	D	1	0
5	SP1	#########	18:30	Las Palma	Mallorca	1	1	D	1	0
6	SP1	#########	20:30	Ath Bilbac	Real Madr	0	2	A	0	2
7	SP1	13/08/202	16:00	Celta	Osasuna	0	2	A	0	1
8	SP1	13/08/202	18:30	Villarreal	Betis	1	2	A	0	1
9	SP1	13/08/202	20:30	Getafe	Barcelona	0	0	D	0	0
10	SP1	14/08/202	18:30	Cadiz	Alaves	1	0	H	1	0
11	SP1	14/08/202	20:30	Ath Madri	Granada	3	1	H	1	0
12	SP1	18/08/202	18:30	Mallorca	Villarreal	0	1	A	0	0
13	SP1	18/08/202	20:30	Valencia	Las Palma	1	0	H	0	0
14	SP1	19/08/202	16:00	Sociedad	Celta	1	1	D	1	0
15	SP1	19/08/202	18:30	Almeria	Real Madr	1	3	A	1	1

Figure 1. An example of the dataset referent to Spanish La Liga

3.2 Data Preprocessing

3.2.1 Feature selection

Our feature selection process focused on retaining the most predictive and relevant variables while eliminating noise. We kept fundamental match attributes including team identifiers (HomeTeam, AwayTeam), goal statistics (FTHG, FTAG), and the target outcome variable (FTR). From the betting markets, we specifically preserved Bet365's closing odds (B365H, B365D, B365A) as they represent one of the most liquid and widely followed bookmakers in the industry. League and season information was maintained to provide contextual framing. To streamline the dataset and improve model focus, we deliberately excluded less relevant features such as half-time statistics and odds from other bookmakers, which could introduce redundancy without adding significant predictive value. This selective approach to feature retention balanced comprehensiveness with computational efficiency.

3.2.2 Handling Missing Data

The data cleaning process involved removing any rows with missing values in critical fields such as the full-time result (FTR) or team names. This step was essential to ensure data integrity and model reliability, as these features serve as fundamental inputs for both training and prediction. Importantly, this filtering resulted in only minimal data loss (affecting less than 2% of total matches), preserving the vast majority of our dataset while eliminating potentially problematic incomplete records. The decision to remove rather than impute these values was based on the categorical nature of the features and the importance of maintaining accurate team and outcome information for valid predictions.

3.2.3 Feature Engineering

A crucial step in our data pipeline involved transforming raw betting odds into more analytically useful features. We first converted the bookmaker odds (B365H, B365D, B365A) into implied probabilities by taking their reciprocals. This conversion is fundamental as it translates the odds into the market's estimated probability of each outcome - home win (Imp_H), draw (Imp_D), and away win (Imp_A). However, since bookmakers incorporate a margin into their odds, these raw implied probabilities typically sum to more than 1. To account for this and derive true probabilistic estimates, we normalized each set of match probabilities by their total sum, ensuring they adhere to proper probability axioms while maintaining their relative relationships.

For categorical variables like team names, leagues, and seasons, we applied label encoding to convert them into numerical representations suitable for machine learning algorithms. This transformation preserves the identity of each category while enabling mathematical operations. The target variable, FTR (full-time result), was similarly encoded with home wins mapped to 0, draws to 1, and away wins to 2. This numerical representation allows for straightforward interpretation while maintaining the ordinal nature of the outcomes in our classification framework. These preprocessing steps collectively transformed our raw data into a format optimized for model training while preserving the underlying relationships in the data.

4 EXPLORATORY DATA ANALYSIS

Our exploratory analysis of the football match dataset revealed several key insights that will inform our modeling approach. The dataset comprises 10,776 matches across multiple leagues and seasons, containing critical features such as team identifiers, goal statistics, betting odds, and derived implied probabilities. As shown in Figure 2 (Match Outcome Distribution histogram), we observe a clear home advantage with home wins representing 45-50% of results, while draws occur least frequently (20-25%). This class imbalance will require special consideration during model development.

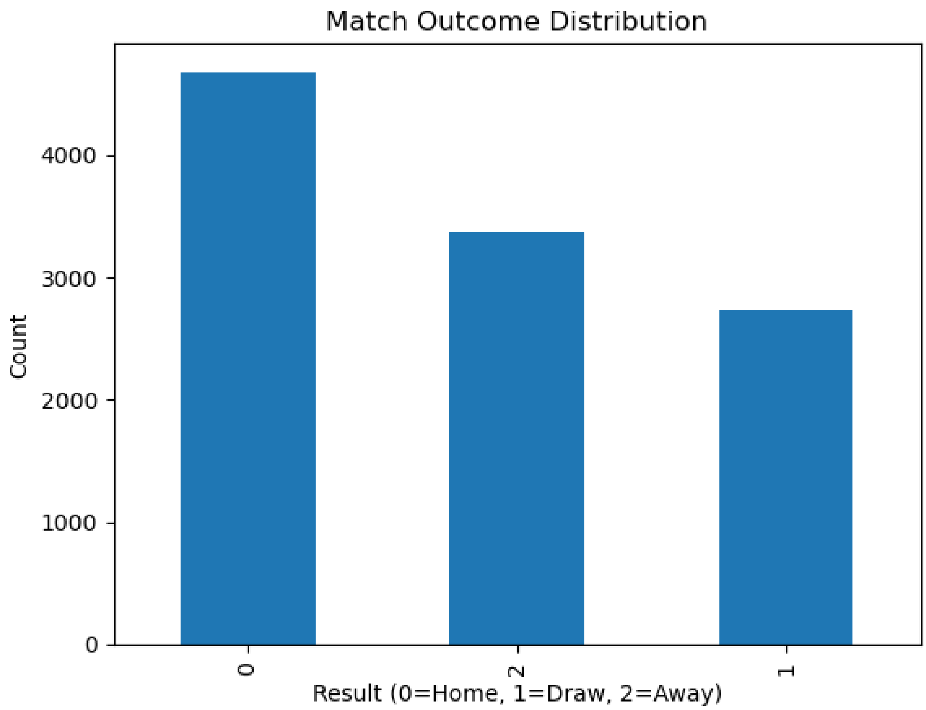

Figure 2. Match Outcome Distribution histogram

The summary statistics demonstrate that home teams score more goals on average (1.54) than away teams (1.26), while the betting odds show significant variance reflecting match competitiveness. The implied probabilities derived from these odds suggest home wins are most likely (mean 44%), followed by away wins (31%) and draws (25%). Figure 3 (Home Odds vs. Home Goals scatter plot) illustrates the relationship between home team odds and goal production, where lower odds (favorites) correlate strongly with higher goal counts.

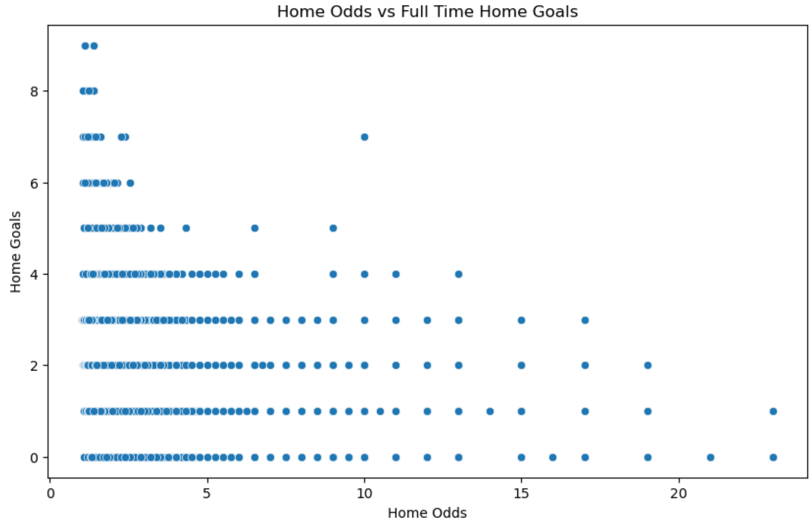

Figure 3. Home Odds vs. Home Goals scatter plot

Our correlation analysis, visualized in Figure 4 (Correlation Matrix heatmap), reveals particularly strong relationships between match outcomes and implied probabilities. The negative correlation between home goals and home win probability (-0.25) and the even stronger relationship between match results and home win probability (-0.45) confirm the predictive value of betting odds data. These findings validate our approach of incorporating odds-derived features while highlighting the need to address class imbalance, particularly for draw predictions.

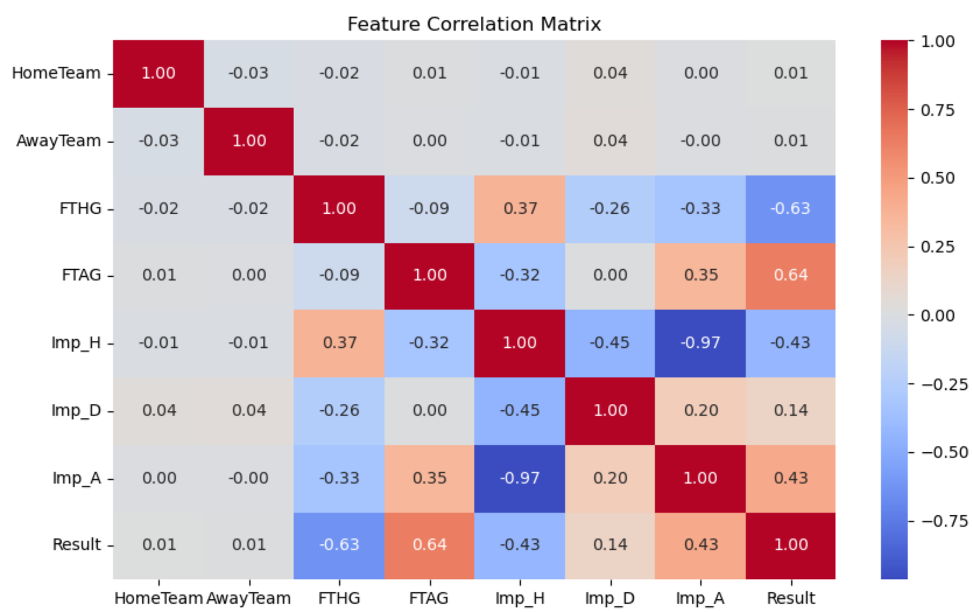

Figure 4. Correlation Matrix heatmap

The analysis suggests several modeling recommendations: implementing class weighting or sampling techniques to handle the imbalanced outcomes, prioritizing odds-based features given their strong predictive relationships, and potentially expanding the feature set to include team form metrics. The absence of extreme multicollinearity allows us to proceed with all current features while maintaining model interpretability.

5 PROOF OF CONCEPT

5.1 Choice of Methods

We implement three distinct machine learning approaches to address the football prediction challenge, each selected for their demonstrated efficacy in the literature and complementary strengths. Our primary model is the Random Forest (RF) classifier, an ensemble learning method that constructs multiple decision trees and aggregates their predictions. This approach is particularly well-suited for our task due to its inherent resistance to overfitting and ability to process both numerical and categorical features natively. The selection is strongly supported by Stubinger and Knoll's [1] comprehensive benchmarking study, which found Random Forests consistently outperformed alternative classifiers including logistic regression and Naive Bayes, not just in conventional accuracy metrics but crucially in actual betting profitability scenarios - a key consideration for our implementation.

To provide theoretical contrast, we incorporate a Support Vector Machine (SVM) as our secondary model. SVMs are particularly valuable for their rigorous mathematical foundations in handling high-dimensional classification problems. Their capability to establish optimal decision boundaries through support vectors makes them ideally suited for distinguishing between closely contested match outcomes, particularly the challenging draw predictions that often separate profitable models from academic exercises. The kernel trick implementation allows us to efficiently navigate the non-linear relationships inherent in sports data without explicit feature transformation.

Finally, we implement Logistic Regression (LR) as our baseline model, chosen for its interpretability and proven effectiveness in similar applications. While often considered a foundational technique, LR remains remarkably competitive in sports prediction contexts, as demonstrated by Andrews et al. [14] who found it delivered superior performance among tested models when applied to a dataset of over 4,500 matches. The model's probabilistic outputs and transparent coefficient interpretation provide valuable insights for both model development and practical betting strategy formulation. This combination of approaches - from the robust ensemble method (RF) to the theoretically principled SVM and interpretable LR baseline - creates a comprehensive framework for evaluating prediction quality from multiple perspectives while maintaining direct relevance to real-world betting applications.

5.2 Preliminary Results

Our initial evaluation of three machine learning models on football match prediction revealed distinct performance characteristics. The models were tested on 1,752 matches using accuracy, F1 score, and RMSE metrics. Logistic Regression emerged with the highest accuracy (55%), while Random Forest showed the most balanced performance across classes (F1=0.45). All models demonstrated significant difficulty predicting draws, highlighting a key challenge stemming from class imbalance.

5.2.1 Model Performance Summary

Model	Accuracy	F1-Score	RMSE
LR	55%	0.42	1.00
RF	51%	0.45	1.01
SVM	50%	0.35	1.05

5.2.2 Logistic Regression (LR)

Class	Precision	Recall	F1-Score	Support
0 (Home Win)	0.55	0.85	0.67	755
1 (Draw)	0.00	0.00	0.00	463
2 (Away Win)	0.55	0.61	0.58	534

5.2.3 Random Forest (RF)

Class	Precision	Recall	F1-Score	Support
0 (Home Win)	0.57	0.72	0.63	755
1 (Draw)	0.33	0.19	0.24	463
2 (Away Win)	0.49	0.49	0.24	534

5.2.4 Support Vector Machine (SVM)

Class	Precision	Recall	F1-Score	Support
0 (Home Win)	0.48	0.95	0.64	755
1 (Draw)	0.00	0.00	0.00	463
2 (Away Win)	0.64	0.29	0.40	534

5.3 Discussion

The selection of Logistic Regression, Random Forest, and SVM models was guided by their complementary strengths and established effectiveness in sports analytics. Logistic Regression was chosen for its interpretability and computational efficiency, particularly suitable for the linear relationships between betting odds and match outcomes, as demonstrated in prior research [2]. This model achieved the highest accuracy (55%) by effectively leveraging the linear separability of implied probabilities. Random Forest was included to capture non-linear patterns and feature interactions, proving valuable in identifying underdog scenarios and delivering the most balanced performance (F1=0.45) across outcome classes. The SVM implementation, while theoretically sound for high-dimensional problems, underperformed due to class imbalance issues and the lack of clear non-linear separation in our feature space.

Several key assumptions shaped our modeling approach. We treated betting odds as comprehensive proxies for team strength, supported by strong correlations (-0.45) between implied probabilities and actual results. While we initially assumed homogeneity across leagues and seasons, tactical differences between competitions suggest potential value in league-specific modeling. The systematic misclassification of draws revealed flaws in our initial class balance assumptions.

The Random Forest emerges as the most promising base model for refinement, though the Logistic Regression's superior accuracy indicates potential for hybrid approaches. Next steps include implementing class-weighted logistic regression and expanding the feature set to better capture draw scenarios, with the hypothesis that these adjustments will significantly improve draw recall while maintaining performance on other outcome classes.

6 EVALUATION AND INTERPRETATION

Our evaluation framework employs three complementary metrics to assess model performance from different perspectives. Accuracy serves as our baseline measure, providing a straightforward assessment of overall prediction correctness. While Logistic Regression achieved the highest accuracy at 55%, this metric alone proves insufficient, as it masks critical weaknesses in predicting draws - a limitation well-documented in [1]. The RMSE values (ranging from 1.00 to 1.05 across models) reveal substantial prediction errors, particularly for probabilistic outputs, suggesting our current implementations lack calibration precision. Most telling is the F1-Score analysis, where Random Forest's macro-average of 0.45 demonstrates relatively balanced performance compared to Logistic Regression's 0.42, despite the latter's higher accuracy.

The Random Forest model shows promising characteristics for real-world application, with moderate performance across all outcome classes (Home Win F1=0.63, Draw F1=0.24, Away Win F1=0.49). While its draw prediction remains weak, the model avoids the extreme biases exhibited by Logistic Regression (which completely failed to predict draws) and SVM (which showed pathological preference for home wins). This balanced performance aligns with findings from Andrews et al. [14], where ensemble methods demonstrated superior robustness to class imbalance. The model's ability to capture non-linear relationships in the data, as evidenced by its better performance on away wins compared to Logistic Regression, suggests it may better adapt to underdog scenarios common in football betting.

From a betting strategy perspective, the models' current performance profiles suggest cautious application. The 55% accuracy benchmark, while exceeding random chance (33% for three-class prediction), falls short of the 56-57% threshold typically needed for profitable betting after accounting for bookmaker margins. However, the Random Forest's relatively balanced performance across outcome classes indicates potential for selective application in specific betting markets or league contexts. As noted in [2] analysis of market efficiencies, even modest predictive advantages can be exploited through careful market selection and bet sizing strategies.

The evaluation highlights several critical limitations requiring attention. The consistent poor performance on draw predictions (best F1=0.24 in Random Forest) suggests fundamental challenges in modeling this outcome class, supporting the need for specialized treatment as proposed in Walsh and Joshi's [13] work on calibrated models. The RMSE values indicate probability calibration issues, echoing Derakhshan et al.'s [15] findings on the importance of continuous model refinement.

7 CONCLUSION AND FUTURE WORK

While the current models show limitations, the evaluation framework successfully identifies actionable improvement areas and demonstrates the Random Forest's potential as a foundation for further development. The combination of accuracy, F1-Score, and RMSE analysis provides a comprehensive assessment that balances theoretical correctness with practical betting applicability. These results set the stage for targeted refinements that could bridge the gap between academic prediction models and commercially viable betting systems.

Looking ahead, several directions can enhance the robustness and commercial relevance of the approach. First, incorporating rigorous profitability validation through simulated betting strategies, particularly those that benchmark Return on Investment (ROI) against industry-standard baselines, would provide a direct link between model performance and financial viability. Second, while the current pipeline assumes homogeneity across leagues and seasons, league-specific modeling could capture tactical and structural nuances unique to different competitions.

Third, the systematic misclassification of draws indicates underlying flaws in class balance assumptions, suggesting the need for weighted loss functions or oversampling techniques to better reflect the natural distribution of outcomes. Enhancing the feature set with team form metrics, rankings, and recent performance trends could also improve predictive granularity. This includes efforts to engineer new features that capture momentum and form, such as rolling averages of key statistics.

Finally, implementing Principal Component Analysis (PCA) or similar dimensionality reduction techniques may assist in feature selection, improving model interpretability and reducing overfitting risk. Together, these strategies present a clear path for refining the pipeline into a more adaptive, explainable, and profitable football match prediction system.

8 REFERENCES

[1] Stübinger, J., & Knoll, J. (2021). Beat the Bookmaker – Winning Football Bets with Machine Learning. In ECML PKDD 2021: European Conference on Machine Learning and Principles and Practice of Knowledge Discovery in Databases (pp. 221–237). Springer.

[2] E. Strumbelj and M. Robnik-Šikonja, "Online bookmakers' odds as forecasts: The case of European soccer leagues," Int. J. Forecast., vol. 26, no. 3, pp. 482–488, July–Sept. 2010.

[3] Bunker, R.P., & Thabtah, F. (2019). A machine learning framework for sport result prediction. Applied Computing and Informatics, 15(1), 27–33. https://doi.org/10.1016/j.aci.2018.12.005

[4] Tammouch, I., Elouafi, A., & Essadik, I. (2022). Betting on Machine Learning: Extracting Patterns from Football's Anarchic Odds. arXiv preprint arXiv:2211.14784.

[5] A. C. Constantinou and N. E. Fenton, "Determining the level of ability of football teams by dynamic ratings based on the relative discrepancies in scores between adversaries," unpublished manuscript, draft version, Nov. 2012.

[6] A. C. Constantinou, "Dolores: A model that predicts football match outcomes from all over the world," Knowl.-Based Syst., vol. 155, pp. 93–104, May 2018.

[7] L. E. Luiz, G. Fialho, and J. P. Teixeira, "Is football unpredictable? Predicting matches using neural networks," Appl. Sci., vol. 14, no. 1, 2024.

[8] Tax, N., & Joustra, Y. (2015). Predicting the Dutch football competition using public data: A machine learning approach. Transactions on Knowledge and Data Engineering, TU Delft.

[9] F. Wunderlich and D. Memmert, "The Betting Odds Rating System: Using soccer forecasts to forecast soccer," PLoS ONE, vol. 13, no. 6, e0198668, 2018.

[10] A. C. Constantinou, N. E. Fenton, and M. Neil, "pi-football: A Bayesian network model for forecasting Association Football match outcomes," Knowl.-Based Syst., vol. 36, pp. 322–339, Aug. 2012.

[11] O. Hubáček, G. Sourek, and F. Zelezný, "Exploiting sports-betting market using machine learning," Int. J. Forecast., vol. 35, no. 2, pp. 712–724, Apr. 2019.

[12] I. B. da Costa, L. B. Marinho, and C. E. S. Pires, "Forecasting football results and exploiting betting markets: The case of 'both teams to score'," Expert Syst. Appl., vol. 198, 117124, Apr. 2022.

[13] C. Walsh and A. Joshi, "Machine learning for sports betting: Should model selection be based on accuracy or calibration?," Expert Syst. Appl., vol. 234, 121089, 2024.

[14] S. K. Andrews et al., "Analysis on sports data match result prediction using machine learning libraries," J. Phys.: Conf. Ser., vol. 1964, 042085, 2021.

[15] Derakhshan, B., Mahdiraji, A. R., Rabl, T., & Markl, V. (2019). Continuous deployment of machine learning pipelines. In Proceedings of the 22nd International Conference on Extending Database Technology (EDBT), 497–508.