Proceedings of CONF-APMM 2026 Symposium: Applied Mathematics and Quantitative Modeling
DOI: 10.54254/2753-8818/2026.PU36424

Mathematical Modeling Strategies and Challenges for Inferring Football Match Results

Shihan Wu

Tsinglan School, Dongguan, China
Shihan.Wu_27@tsinglan.org

Abstract. This paper reviews mathematical and computational prediction models which forecast football match outcomes. It also summarizes common models, data inputs, and key challenges. Based on a survey of statistical methods including Poisson and Elo, machine learning, neural network, and deep-learning approaches, the review compares how different techniques use match-level, team-level, player-level and external features. It specifically shows which inputs tend to improve prediction, such as in-game statistics and player information, as well, and it explains typical evaluation practices and performance limits. Major obstacles include noisy or incomplete data, the inherent randomness of sport, class imbalance especially matches with draw outcomes, and the computational cost of complex models. The paper also points out practical directions to improve reliability, including combining diverse datasets, and focusing on model interpretability so coaches and analysts can trust outputs and apply them to improve team performances. Overall, although many methods show a good future potential, significant doubts still remain. In future works, the priority should be on gathering more reliable data and improving how to explain the models.

Keywords: football prediction, prediction models, machine learning, sports analytics, data analysis

1. Introduction

A huge number of people around the world are playing football and it's also regarded as one of the world's most popular sports. According to a survey, there are almost five billion football fans nowadays in the world [1]. What's more, 2022 FIFA World Cup attracted more than 1.5 billion viewers. This huge popularity means that football match outcomes are very important to not only most of the fans, but also to clubs and bettors. Therefore, prediction of football match results has various precious value. Firstly, football clubs use outcome forecasts to refine tactics. Furthermore, analysts study performance indicators to guide coaching decisions. Betting markets depend on forecasts to set odds as well. Football has become very data-driven as prediction of football match results provides clubs with valuable insights and a competitive advantage. Moreover, experienced bettors can use their own trained prediction model to predict games for profit. Likewise, coaches of teams may use statistical models to identify key performance indicators. For example, Parim et al. [2] find that scoring the first goal can significantly increase shots on target and significantly increase a team's chance of winning, which may help coaches develop different strategies based on game

© 2026 The Authors. This is an open access article distributed under the terms of the Creative Commons Attribution License 4.0 (https://creativecommons.org/licenses/by/4.0/).

DOI: 10.54254/2753-8818/2026.PU36424

context. In sum, reliable forecasting of football match results is important across betting, tactical analysis, and sports media contexts.

These years, data analysis and forecasting techniques have become the core of sports analytics. Machine learning and statistic models allow researchers to find out the potential patterns from historical match data and real-time statistics. Some more modern methods like neural networks are also been applied to predict football match outcome by more and more researchers though these methods are not as stable and well-developed as the classic ways. Using different features from past data and real-time statistics, lots of satisfying prediction models can give a result for reference and have a considerable accuracy. In addition, these models provide football matches with a new dimension, which is for coaches to transfer a team's arrangement for a competition with help of those models' outcome prediction.

This paper aims to summarize and synthesize the main approaches to football match prediction, the types of data they employ, and the core challenges they face, with surveys of the literature on mathematical and computational models used for forecasting match outcomes from traditional statistical models to modern machine-learning algorithms. What's more, there is a catalog of the typical data inputs, including team standings, player stats, situational factors, betting odds, etc. This paper also discuss the inherent difficulties highlighted in the literature, such as data quality, the random nature of football, and model limitations, that complicate accurate forecasting. By organizing this information, the paper reviews the stage for understanding current capabilities and gaps in football match prediction methods.

2. Data collection and analysis

2.1. Data source

There are various open sources of datasets for training a football prediction model on all kinds of websites.

Common choices include official match statistics, such as the league and data providers (e.g. Opta), public datasets from websites like Kaggle, and betting market data. Furthermore, some research even uses data collected from football match simulation game with predefined match conditions [3]. For example, Calvin C. K. Yeung et al. [4] point out the importance of open source datasets, and use a dataset "European Soccer Database" from Kaggle, which contains detailed match records including team formations, players selected, and the player ratings for each match. Moreover, they also use data from European Soccer Supplementary Database as a complementary source.

Betting odds are also widely used as a significant factor in training of prediction models, as they are set after aggregating expert and market information. Indeed, Yeung et al. [4] notice that weather conditions, expert opinions are also frequently adopted as features in forecasting models.

In addition, there are also some brand new and creative ways of collecting data for training. For instance, Kampakis et al. [5] use Twitter as an input to predict the result of a football match. However, it is still necessary to consider whether it is proper to use this kind of source because there may be lots of noise in the datasets, greatly affecting the accuracy of models.

2.2. Analysis of data

Analyses typically categorize features by different levels. Match-level features are information about the match itself. This level usually includes "Home/Away", "Final Scores", "Shots", etc. As people

DOI: 10.54254/2753-8818/2026.PU36424

know, home advantage can influence the outcome of a football match to a great extent. Sometimes "Attendance" and "Referees" are also considered as match-level variables in models. Team-level features are composed of the condition of clubs, including "Historical Performance", "League Ranking", "Recent Form", etc. For example ,Hvattum et al. [6] talk about that team historical rankings and team ratings such as Elo are usually used in previous prediction models. The third level is player-level. Features in this level are the characteristics or other information of players themselves, such as "Lineup", "Age", "Playing Time", "Injuries", "Players' skill ratings", etc. A great example of using these features is the model from Yeung et al. [4]. When training the model, their feature set includes selected players, formations, and player qualities (FIFA ratings), which can be both considered as player-level features.

Finally, external factors are also considered. Weather condition is a classic external factor. Temperature, precipitation, wind are common weather conditions that might influence results of matches. A previous study trains their models with datasets with detailed weather information such as temperature, wind speed, etc., and the results show that several weather features can considerably improve the performance of the model prediction [7]. Additionally, some other factors like crowd sentiment may also affect a match.

2.3. Importance of different features

Existing research has already identified which features most affect the prediction outcome. Wheatcroft [8] reported that "Shots on or off target" is a strong predictor. In fact, prediction models that first predict the result with non-rare match statistics, such as shots, crosses, etc., and then use those to predict the result achieve substantially better performance than models using only goals. Moreover, models that focus on players themselves may be more accurate than models using data of teams and matches. Yeung et al. [4] compared their model, which uses selected players, formations, and player qualities (FIFA ratings) as major factors, with betting-odds baseline, and they point out that their prediction model achieves higher F1-score (0.47) than the odds baseline (0.39). Features like weather condition or referees typically have a smaller influence, but sometimes they still have non-negligible effects on the outcome.

3. Prediction models for football matches

3.1. Statistical models

The simplest way to predict the outcome of a football match is using the statistical model. This method is widely applied in the early research. One very classic statistical model is the logistic regression. Logistic regression model gives the probability p that the input x belongs to class 1 by passing the linear score z through a "Sigmoid" function shown in formula (1), then learns the weights w that maximize the likelihood of the observed labels. In prediction of football competitions, logistic regression models the probability of 3 outcomes, including win, draw, and lose, into a logit function of input features. Logistic regression model is very easy to build and can provide clear coefficient for each factor. It is very efficient because of its simple principle as well. However, logistic regression can only handle linearly separable problems, which means it assumes a linear effect between factors and outcome of matches. Therefore, logistic regression cannot find out the complex and non-linear relationships between features. As a result, though logistic regression is not a proper choice for predicting football match result compared with other more complex and flexible methods, it still can serve as a baseline generally.

DOI: 10.54254/2753-8818/2026.PU36424

$$\mathrm{f}(z) = \frac{1}{1+e^{-z}} \tag{1}$$

In order to model scores more directly, models based on Poisson come to more researchers' eyes. Maher's [9] classic model considers each team's number of goals as an independent Poisson random variable. The method provides a distribution of probability for each possible scoreline. What's more, it also generate a rating for "Attack" and "Defense" for each team. Dixon and Coles [10] base on the previous model and change it to allow more types of data. The method can greatly ignore the fluctuations of a team's performance during a time period. The updates improve the stability of the model and allow it to perform better under other various conditions. These Poisson models are powerful because they are mostly based on the statistic data of goals and as a result can fit historical match totals. However, Poisson models assume the goal numbers are independent. They need a huge amount of data for training a stable model as well. Furthermore, Poisson models tend to give a prediction that overestimates the score.

Another statistical method of football match outcome is the Elo rating system. This kind of method originate from chess and is extended to a variety of team sports. Elo method gives every team a single rating that is updated after every match. The rating will increase when a team wins and decrease when a team loses, which means that this rating can greatly reflect the strength of a football team. As a result, the result of a match is modeled to a function of ratings of two groups. A fantastic example for using Elo rating system to predict the outcome of a competition is the model by Hvattum and Arntzen [6]. They showed how Elo-based ratings can be used as covariates in a logistic or ordered logit model. Moreover, Havettum and Arntzen [11] introduced two covariates, Elo ratings and individual ratings, and further construct 2 models, a ordered logit regression (OLR) model and a competing risk model. Advantages for rating methods are that these kinds of methods are very simple and perform well in finding the long term strength of a team. Nonetheless, Elo ratings depend on past results very much, so it may be not accurate when facing new team or players. For instance, Yeung et al. [4] notice that although Elo ratings are commonly used in the history, it limits the models from generalization because they assume static, past data.

Statistical models also include Bayesian networks. Constantinou's [12] work mixes the Bayesian networks with dynamic ratings to forecast the outcome of football match and has a positive result. All in all, classical methods like logistic regression, Poisson, Elo ratings, etc., have the strength of transparency and simplicity, but they may not be that proper for complex patterns in modern datasets.

3.2. Machine learning

With the increasing complexity of data nowadays, researchers have applied machine learning methods to train a prediction model not only for football matches but also lots of other fields. Machine learning includes a huge range of models, such as decision trees, random forests, gradient-boosted trees (e.g. XGBoost), support vector machines (SVMs), k-nearest neighbors, and neural nets. From Groll [13] and Rosli's [14] research, it is clear that XGBoost, decision tree, and a lot of other machine learning models are widely used for football match outcome prediction.

Machine learning models based on trees are particular popular among researchers. Decision tree recursively splits the feature space on the attribute that best separates the classes, forming a tree of axis-aligned "if-then" rules, and the prediction is the majority class of the leaf reached. However, a single decision tree often overfits and is not stable. As a result, s random forest contains a number of decision trees and does a majority vote for the given input to determine the final prediction. This action can greatly improve the model stability and accuracy. Gradient boosting machines such as

DOI: 10.54254/2753-8818/2026.PU36424

XGBoost is another type of prediction model in machine learning. This method will sequentially train the trees to fit the features and correct errors. According to Sleeuwaert's [15] work, the method of gradient boost model usually performs the best among various models in football match prediction. Reason for this result may be the strength of gradient boosting in handling heterogeneous features, capturing nonlinear interactions, and providing feature-importance scores. However, in real practice, other methods can also provide a satisfying result. For instance, Rosli [14] discover that decision tree perform very well with an accuracy of 99.56% in the comparison between different machine learning models. Additionally, SVMs have been applied for football matches prediction as well. They can find complex decision boundaries, but they also require careful and proper setting of the coefficients and scale when facing very large datasets.

3.3. Deep learning and neural networks

In recent years, deep learning and neural networks methods are being applied by more and more researchers to predict the result of a football match. Recurrent neural networks (RNNs) and other methods which are based on RNNs, such as Long Short-Term Memory (LSTM), are very appropriate for time series problems, so can be applied for prediction of football matches with related time series features. These models can study the team performance in a time series. For example, Zhang et al. [16] creatively mixed the LSTM model and the attention mechanism. Based on this, they then further suggest an AS-LSTM model for predicting match results, with the addition of a time sliding window to ensure the timeliness of the prediction result. In another application of deep learning models, Rahman [17] builds a deep neural network(DNN) architecture, which successfully predicted 63.3% of the 2018 FIFA world cup matches. Other deep learning methods such as convolutional neural networks (CNNs) and transformer architectures also have been widely used. Hsu [18] applies CNNs and other models to sports event prediction.

Deep learning methods are very powerful in prediction area. However, they are still developing and face a lot of challenges. For instance, the process for deep learning models to predict is very hard to interpret.

4. Discussion

4.1. Model performance and evaluation

Determining whether a prediction model performs well or not needs carefully designed metrics. F1 score, confusion metrics, accuracy, precision, and recall are evaluation metrics that are commonly used. These indicators can be used to compare the performance between different models or to evaluate the improvement of a model. Yeung et al. [4] use F1 score to test their models, as an example, and it turns out that F1 score of their model reaches 0.47, which is higher than the betting-odds baseline(0.37). There are a variety of loss functions being used during training. Classification models use cross-entropy or log-loss for example, while regression models might use mean-squared error. But at last, researchers often present a set of metrics combining a lot of indicators to give a complete report of performance.

Furthermore, the target of the prediction also influences the performance to a great extent. Forecasting the outcome of football matches, which includes win, draw, or lose, is far easier than predicting the exact scores of teams, like 1-1 or 2-3. Therefore, the main goal for most prediction models is to forecast the winner. But, although predicting the exact score is really difficult, it's still meaningful for researchers to discover the way to overcome as the goal number is also a key factor of a football match.

4.2. Data quality and computation cost

Every prediction model needs a huge amount of data to train on, so the quality of the data becomes really significant. Soni et al. [19] suggest that a great statistic can contribute to a wonderful model a lot. Galekwa et al. [20] also mention the importance of data quality. They believe that the data availability is a big issue and some datasets can be incomplete. The risk of "bias" also exists, which means a model may too much rely on historical data, but they only reflect past conditions that may shift over time. Overfitting to the past data of one football league can harm generalization. What's more, it's possible for models which rely too much on the past data to perform well in performance tests as they may also be based on historical matches, but actually it behaves poorly in future real practice.

In addition, the computation cost of models is another important problem. Simple models such as statistical models usually are extremely fast, but they are too simple for complex prediction task. However, complex models including machine learning and deep learning can be slow, and so they are difficult for application of real-time prediction during football competitions. In fact, models must be lightweight enough to recompute probabilities quickly after each new event in real practice, and they may also only have the access to part of the match data. So currently real-time prediction for football matches is still a tough work for recent researchers.

Apart from time requirements, Handling very large datasets can bring stress to the storage and the efficiency of the devices.

5. Conclusion

This paper reviews how mathematical or computational models are used to forecast the result of football matches, and highlights their new discoveries as well as limitations. Studies collectively show that a variety of modeling strategies, ranging from Poisson and Elo-based statistical approaches to machine-learning classifiers and deep learning models, can capture key indicators of team performance and achieve good and predictive accuracy. However, a very serious issue is that football matches include lots of random factors which affect forecasting accuracy to a really great extent. In practice, common challenges such as home-field advantage have also been noticed by researchers. For example, many matches still favor the home team, leading to class imbalances in prediction datasets. These insights suggest that although current models can identify some useful trends, they cannot fully get rid of the uncertainty, especially for rare outcomes like draws. By combining together these findings, this paper offers a review of existing prediction techniques, the data they currently use, and the performance boundaries of models.

Looking ahead, the findings suggest clear directions for future further research. There is growing potential to integrate better and more real-time data sources, such as player tracking, physiological sensors, and advanced event data, into predictive frameworks. Future models might leverage deep learning on tracking data or incorporate social and contextual signals to improve reliability. On the methodological side, techniques such as ensemble learning and rigorous validation, which means combining random forests and boosting with cross-validation as an example, can decrease the effect of overfitting and enhance model robustness. Nevertheless, challenges still exist. Even the best models must deal with incomplete information and shifting team dynamics. Therefore, ongoing work should emphasize not only improving raw accuracy but also enhancing interpretability so that

DOI: 10.54254/2753-8818/2026.PU36424

coaches and analysts can understand model reasoning and address data limitations. By outlining the current landscape and pointing out these gaps, this paper can contribute to development of more effective football match prediction models and help future efforts in sports analytics.

References

[1] FIFA Publications. (2021). *The football landscape: FIFA vision report*. FIFA Publications. https://publications.fifa.com/en/vision-report-2021/the-football-landscape

[2] Parim, C., et al. (2021, July 28). Prediction of match outcomes with multivariate statistical methods for the group stage in the UEFA Champions League. *Journal of Human Kinetics, 79*, 197–209.

[3] Prasetio, D. (2016). Predicting football match results with logistic regression. In *2016 International Conference on Advanced Informatics: Concepts, Theory and Application (ICAICTA)*. IEEE.

[4] Yeung, C. C. K., Bunker, R., & Fujii, K. (2023). A framework of interpretable match results prediction in football with FIFA ratings and team formation. *PLOS ONE, 18*(4), e0284318.

[5] Kampakis, S., & Adamides, A. (2014). Using Twitter to predict football outcomes.

[6] Hvattum, L. M., & Arntzen, H. (2010). Using ELO ratings for match result prediction in association football. *International Journal of Forecasting, 26*(3), 460-470.

[7] Fan, Y., & Knobbe, A. (2023). Predicting football matches with the use of weather conditions. *Leiden Institute of Advanced Computer Science (LIACS)*.

[8] Wheatcroft, E. (2021). Forecasting football matches by predicting match statistics. *Journal of Sports Analytics, 7*(2), 77-97.

[9] Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica, 36*(3), 109-118.

[10] Dixon, M. J., & Coles, S. G. (1997). Modelling association football scores and inefficiencies in the football betting market. *Journal of the Royal Statistical Society: Series C (Applied Statistics), 46*(2), 265-280.

[11] Arntzen, H., & Hvattum, L. M. (2021). Predicting match outcomes in association football using team ratings and player ratings. *Statistical Modelling, 21*(5), 449-470.

[12] Constantinou, A. C. (2019). Dolores: A model that predicts football match outcomes from all over the world. *Machine Learning, 108*(1), 49-75.

[13] Groll, A., et al. (2018). Prediction of the FIFA World Cup 2018 - A random forest approach with an emphasis on estimated team ability parameters.

[14] Rosli, C. M. F. C. M., et al. (2018). A comparative study of data mining techniques on football match prediction. *Journal of Physics: Conference Series, 1020*(1). IOP Publishing.

[15] Sleeuwaert, T. (2020). *Evaluation of the current state of football match outcome prediction models* (Master's dissertation). Ghent University.

[16] Zhang, Q., et al. (2022). Sports match prediction model for training and exercise using attention-based LSTM network. *Digital Communications and Networks, 8*(4), 508-515.

[17] Rahman, M. A. (2020). A deep learning framework for football match prediction. *SN Applied Sciences, 2*(2), 165.

[18] Hsu, Y.-C. (2021). Using convolutional neural network and candlestick representation to predict sports match outcomes. *Applied Sciences, 11*(14), 6594.

[19] Soni, A., et al. (2023). Evaluating the impact of data quality on machine learning model performance. *J. Nonlinear Anal. Optim, 14*(1), 13-18.

[20] Galekwa, R. M., et al. (2024). A systematic review of machine learning in sports betting: Techniques, challenges, and future directions.