Proceedings of ICFTBA 2026 Symposium: Strategic Management & Business Analytics: Data-Driven Performance Models
DOI: 10.54254/2754-1169/2026.LD39176

The Effectiveness and Limitations of Artificial Intelligence in Predicting FIFA World Cup Match Outcomes

Yafei Zhang

International Department, Yinchuan No.2 Middle School, Yinchuan, China
Zhangyafei0918@gmail.com

Abstract. Artificial intelligence (AI) is being used for prediction in many areas, including finance, healthcare, and sports. Football is a harder case because a match can change quickly and the result is not decided by statistics alone. The FIFA World Cup is a useful example because teams from different regions meet in a short tournament, and small changes in form or tactics can matter. This paper looks at how AI methods have been used to predict FIFA World Cup match results. It reviews several common approaches, including Logistic Regression, Random Forest, Support Vector Machine (SVM), Artificial Neural Networks (ANN), and Transformer models. The paper compares these methods in terms of prediction ability, interpretability, data needs, and practical use. AI has some clear advantages. It can work with large amounts of data and find patterns that may be difficult to see with simple statistical methods. At the same time, its predictions are limited by the quality of the data and by factors that are hard to measure. Player injuries, referee decisions, tactical changes, weather, and psychological pressure can all affect a match. These factors may appear suddenly and may not be included well in historical datasets. The main conclusion is that AI can be useful for football prediction, but it should not be treated as a perfect answer. It is more useful as a tool that supports human judgment. This study compares existing research and suggests several ways that future football prediction models could be improved.

Keywords: Artificial Intelligence, Machine Learning, FIFA World Cup, Sports Analytics, Match Outcome Prediction, Predictive Modeling

1. Introduction

Artificial intelligence (AI) and machine learning (ML) are now common topics in sports analytics. Football prediction is one area where they are often used. Predicting the outcome of a World Cup match is not simple because the result can be affected by team strength, tactics, weather, referee decisions, and the performance of individual players [1, 2].

Earlier prediction methods mainly used historical measures such as FIFA rankings and head-to-head records, often within basic statistical models, but these measures did not always capture non-linear relationships [3, 4]. More recent studies have used machine learning algorithms (e.g., Logistic Regression, Random Forest, Support Vector Machines, Neural Networks, and Transformers) to identify patterns in larger datasets and estimate outcome probabilities [5-10]. AI models still face

© 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: 10.54254/2754-1169/2026.LD39176

problems with data quality and model interpretability ("black-box" nature), as well as difficulties in accounting for unpredictable factors such as player morale, tactical surprises, or mid-match red cards [2, 11]. An important practical question is therefore whether AI should be treated as a final predictor or as a tool that supports human judgment.

This study uses a literature-based comparison and proposes a conceptual framework for FIFA World Cup outcome prediction. To achieve these objectives, this study addresses the following research questions: (1) How effectively can artificial intelligence algorithms predict FIFA World Cup match outcomes? (2) Which machine learning algorithms demonstrate the greatest advantages and limitations in football prediction? (3) What factors prevent artificial intelligence from achieving consistently accurate predictions in highly uncertain sporting events such as the FIFA World Cup?

2. Literature and theory

Sports analytics has changed quite a lot over time. Early football prediction studies mainly used statistical models to estimate the number of goals or the chance of different results. Xu and Zeng discussed the mathematical properties of the Poisson distribution [12]. Later, Maher and Dixon and Coles applied related ideas to football scores [3, 4]. These methods are relatively easy to understand, but their assumptions can make it difficult to describe complicated tactical situations.

Machine learning was introduced partly because football data contains many variables. Instead of setting every relationship in advance, ML models can learn patterns from the data [7]. Some early sports studies also combined regression with machine learning. Luo et al. used these methods to predict Olympic medals, while Wu used an Extreme Learning Machine to reduce errors in a traditional grey forecasting model [6, 13, 14].

Football research has also started to compare several machine learning models directly. Song et al. studied the 2022 FIFA World Cup with Logistic Regression, Decision Trees, Random Forests, Support Vector Machines, AdaBoost, and Artificial Neural Networks [5]. Their results showed that some ensemble methods and neural networks could perform better than basic statistical models when the input features were carefully selected [2, 5, 8, 10].

Deep learning has added another level of complexity. Horton and Lucey proposed an Axial Transformer Neural Network for predicting match results, team performance, and player statistics [6]. The model can use information from different points in a sequence, which is useful when match events change over time. However, this type of model needs more computing power and more data.

The literature also points to several gaps. Model results depend strongly on the features used and the quality of historical data. Historical data may also favor teams that have traditionally been strong. Some important factors, such as psychological pressure, a sudden injury, or a referee's decision, are difficult to measure. Deep learning models also have limited interpretability [11]. These problems make it useful to compare different models using the same basic criteria.

3. Method

3.1. Research design

This study uses a qualitative comparative analysis of existing research. It does not train a new machine learning model with a new historical match dataset. Instead, it examines previous studies, compares several representative prediction methods, and organizes their main features into a conceptual framework.

DOI: 10.54254/2754-1169/2026.LD39176

The comparison focuses on four practical questions. First, how well can each method deal with football data and non-linear relationships? Second, how easy is it to understand why a model gives a certain result? Third, how much data and training does the model need? Finally, how practical is the method when it is used for football analysis? These points are used throughout the comparison rather than treating prediction accuracy as the only measure of a useful model.

3.2. Feature categories

The prediction features identified in the football studies are divided into two groups based on the football prediction studies reviewed in this paper [3, 5].

(1) Long-Term Team Strength Indicators: These include FIFA World Rankings, World Football Elo Ratings, total squad market value, and previous head-to-head (H2H) results. They give a general picture of the strength of each national team.

(2) Recent Performance and Form Indicators: These include recent win rates, recent scoring performance, defensive statistics such as goals conceded and clean sheets, and results from official international qualifiers played before the tournament.

3.3. Analytical evaluation dimensions

The models are compared from four main angles:

(1) Predictive Capability: whether the model can produce useful predictions, deal with non-linear relationships, and work with many variables.

(2) Interpretability: how clearly a user can understand the reason behind the model's prediction.

(3) Data Requirements: how much data the model needs and how strongly its results depend on data quality and feature selection.

(4) Practical Applicability: how easy the model is to use, how much computing power it needs, and whether it can provide results quickly.

4. Analysis

4.1. Comparison between traditional statistical models and machine learning approaches

Traditional statistical methods, including Poisson regression and Logistic Regression, are easier to explain because their mathematical structure is clear [3, 4, 12]. They also do not usually need as much training data and can run quickly. Their weakness is that simple assumptions may not describe all the relationships in a football match. For example, the effect of midfield control may depend on the opponent's counter-attacking style, which is difficult to represent with a simple linear relationship.

Machine learning methods can deal with more complicated relationships [7]. They can combine information about players, teams, tactics, and recent results. In several football prediction studies, this has led to better results than basic statistical models [5]. However, better prediction does not automatically mean that the model is more useful. More complex models can require more data and may be harder to explain.

4.2. Performance and applicability of machine learning algorithms

A comparative synthesis of widely applied machine learning algorithms in football analytics is summarized in Table 1:

DOI: 10.54254/2754-1169/2026.LD39176

Table 1. Comparative evaluation of algorithms for FIFA World Cup prediction

Algorithm	Strengths	Limitations	Football Application Context
Poisson Distribution /Logistic Regression [3, 4, 12]	High interpretability, low data and computational requirements	Assumes linear relationships, fails to capture complex non-linear dynamics	Baseline win/loss/draw outcome and scoreline probability estimation
Random Forest [8]	Excellent non-linear modeling, resistant to overfitting, provides feature importance ranking	Can become computationally heavy on massive trees, limited extrapolation capability	General match result classification based on multi-dimensional statistics [5]
Support Vector Machine [9]	Highly effective in high-dimensional spaces, robust against small sample noise	Sensitive to parameter tuning and kernel selection, poor scalability to massive data	Binary-Multiclass outcome classification in small-sample tournament settings [5]
Artificial Neural Networks [2]	Strong capability in learning intricate, non-linear feature representations	High risk of overfitting on small datasets, low interpretability ("black box")	Modeling non-linear tactical interactions and complex team performance
Transformer Models [6]	Superior handling of sequential time-series data and dynamic in-game events	Demands massive real-time datasets, high computational cost and optimization complexity	Real-time, in-game outcome forecasting and player-level dynamic tracking [6]

4.3. Deep learning models and emerging AI technologies

Transformer-based models are a newer option for football prediction [6, 11]. Their main advantage is that they can work with information that comes in a sequence. In a football match, events happen one after another, so a model can use earlier events when estimating what may happen next.

This approach may also allow a model to combine different kinds of information. Future systems may combine numerical match statistics together with text, such as injury news, pre-match interviews, or reports about tactics [1]. However, these models are more demanding than simpler methods, both in data and computing resources.

4.4. Limitations and uncertainty of AI-based football prediction

AI prediction has several important limits in World Cup matches:

(1) Unquantifiable Qualitative Variables: Some important events are difficult to turn into reliable numerical variables. A sudden injury, pressure during a penalty shootout, a red card, or bad weather can change a match quickly. Historical training data cannot fully represent these situations.

(2) Algorithmic Bias and Historical Overfitting: A model trained mainly on historical results may give too much weight to teams that have been strong for many years. This can make it harder for the model to recognize a team that has improved quickly. Morocco's performance at the 2022 World Cup illustrates how rapid changes in team performance may be difficult for models that rely heavily on older historical data to capture.

(3) Black-Box Nature: More complex models, including ANNs and Transformers, can be difficult to explain [2, 11]. A high prediction score is less useful if users cannot understand which factors led to the result.

DOI: 10.54254/2754-1169/2026.LD39176

5. Recommendation

The comparison in this study suggests several possible directions for future work. These recommendations mainly focus on improving data quality, model performance, interpretability, and practical use.

Use More Types of Data: Future models could use more than rankings and goal statistics. Player tracking, workload, weather, and referee-related information may provide useful additional signals.

Combine Different Models: A simple statistical model such as a Poisson model can provide an understandable baseline, while a model such as Random Forest or XGBoost can handle more complicated relationships. Combining these approaches may provide a useful way to balance predictive performance and interpretability.

Make Models Easier to Explain: Methods such as SHAP and LIME could be used to identify which features contribute most to a prediction. This would make the results easier for coaches and analysts to examine.

Include Text and Other Unstructured Information: Future systems could use information from injury reports, news, and coaching press conferences. Large Language Models may help turn this information into variables that can be used alongside normal match statistics.

Use AI as Support, Not as the Final Answer: Football contains too many uncertain events for an AI model to give a guaranteed result. AI predictions may be more useful when considered together with human knowledge of teams, players, tactics, and the situation of a match.

6. Conclusion

This paper compared traditional statistical methods, machine learning models, and newer deep learning approaches for FIFA World Cup match prediction. The studies reviewed in the paper suggest that methods such as Random Forest, Neural Networks, and Transformers can perform well in some football prediction tasks. Their advantage is mainly their ability to work with more variables and capture more complicated relationships.

At the same time, these models have clear limits. Their results depend on the data used to train them, and historical data can contain bias toward teams that have been successful for a long time. More importantly, some match factors are difficult to predict before they happen. Player morale, tactical changes, referee decisions, injuries, and other unexpected events can all affect the final result.

AI can therefore be useful for World Cup analysis, but it should not replace human judgment. Future work could improve prediction by using more up-to-date data, combining different model types, and making model decisions easier to understand. Even with these improvements, the uncertainty of football means that a prediction will remain a probability rather than a guarantee.

However, AI models remain fundamentally constrained by data quality, algorithmic bias toward historically dominant teams, and the inability to quantify dynamic, unpredictable qualitative variables such as player morale, tactical surprises, referee decisions, or mid-game injuries. AI should therefore support human judgment rather than replace it. Future advances in Explainable AI, multimodal learning, and real-time data integration will continue to enhance sports analytics, while the intrinsic unpredictability of the FIFA World Cup will ensure football remains a compelling, human-driven contest.

DOI: 10.54254/2754-1169/2026.LD39176

References

[1] Russell, S., & Norvig, P. (2021). Artificial intelligence: A modern approach (4th ed.). Pearson.
[2] Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep learning*. MIT Press.
[3] 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.
[4] Maher, M. J. (1982). Modelling association football scores. *Statistica Neerlandica, 36*(3), 109–118.
[5] Song, Y., Sun, G., Wu, C., Pang, B., Zhao, W., & Zhou, R. (2024). Construction of 2022 Qatar World Cup match result prediction model and analysis of performance indicators. *Frontiers in Sports and Active Living, 6*, 1410632. https://doi.org/10.3389/fspor.2024.1410632
[6] Horton, M., & Lucey, P. (2025). Large-scale in-game outcome forecasting for match, team and players in football using an axial transformer neural network. arXiv preprint arXiv:2511.18730.
[7] Bishop, C. M. (2006). *Pattern recognition and machine learning*. Springer.
[8] Breiman, L. (2001). Random forests. *Machine Learning, 45*(1), 5–32.
[9] Cortes, C., & Vapnik, V. (1995). Support-vector networks. *Machine Learning, 20*(3), 273–297.
[10] Friedman, J. H. (2001). Greedy function approximation: A gradient boosting machine. *Annals of Statistics, 29*(5), 1189–1232.
[11] LeCun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature, 521*(7553), 436–444.
[12] Xu, Y., & Zeng, M. (2006). Research on the properties and applications of the Poisson distribution. *Journal of Yangtze University (Natural Science Edition), 3*(4), 69–72.
[13] Luo, W., Zhou, B., & Tian, X. (2025). An Olympic medal prediction model based on regression and machine learning. *Tianfu Mathematics, 2025*(7).
[14] Wu, P. (2017). A sports performance prediction model using an extreme learning machine to correct prediction errors. *Modern Electronics Technique, 40*(3), 117–120.