World Journal of Information Technology
Print ISSN: 2959-9903
Online ISSN: 2959-9911
DOI: https://doi.org/10.61784/wjit3123

A MULTI-MODEL ENSEMBLE LEARNING FRAMEWORK FOR COMPLEX EVENT FORECASTING

MengYun Shi[1#,*], YaLan Jin[2#]
[1]*School of Information and Communication Engineering, Communication University of China, Beijing 100024, China.*
[2]*The School of Data Science and Intelligent Media, Communication University of China, Beijing 100024, China.*
[#]*These authors contributed equally.*
[*]*Corresponding Author: MengYun Shi*

Abstract: Forecasting outcomes in complex and multi-factor environments remains a significant challenge due to nonlinear interactions, temporal dependencies, and high-dimensional variables. This study presents a robust multi-model ensemble framework designed for large-scale performance prediction in dynamic competitive scenarios. The approach integrates XGBoost for nonlinear feature interactions, LSTM for temporal sequence modeling, and ElasticNet for interpretable linear correlations, combined through a stacking ensemble strategy to improve predictive accuracy and robustness. Furthermore, the framework incorporates uncertainty quantification using confidence intervals, enabling risk-aware decision-making. Experimental validation demonstrates that the proposed ensemble model outperforms individual models across multiple evaluation metrics and provides interpretable, actionable insights for strategic resource planning and performance assessment.
Keywords: Forecast; Ensemble learning; XGBoost; LSTM; ElasticNet

1 INTRODUCTION

Complex competitive environments often involve diverse influencing factors, including historical performance, participation structure, and event diversity, which exhibit strong nonlinear and temporal dependencies. Traditional single-model approaches struggle to accurately capture these intricate relationships, leading to limited predictive reliability in large-scale performance forecasting tasks.

With the rapid growth of data availability and computational resources, machine learning has emerged as a powerful tool for addressing such multi-dimensional prediction problems. In particular, ensemble learning, which combines complementary models, can effectively handle multi-source and high-dimensional data, improving both accuracy and robustness. Moreover, uncertainty quantification and model interpretability are increasingly important for providing actionable insights to decision-makers. Motivated by these needs, this study proposes a Stacking ensemble framework integrating XGBoost, LSTM, and ElasticNet to enhance prediction accuracy while offering interpretable results and confidence intervals.

Medal outcomes are influenced by population, economic strength, event diversity, and host-nation advantages, with complex nonlinear and temporal dependencies[1]. Early studies mainly relied on econometric and time-series models, revealing significant correlations between population size, GDP per capita, and medal counts[2-3], while also highlighting the role of geography, cultural traditions, and talent migration in short-term breakthroughs[4]. Host-nation effects have been consistently validated, showing a clear boost to both total and gold medal counts, though the impact varies by edition and event [5-6].

With the growth of data and computing power, machine learning has become the mainstream approach. Neural networks were initially used to capture complex nonlinear relationships but were prone to overfitting [7-8]. Recently, deep learning and ensemble methods have significantly improved accuracy, with models like random forests, SVMs, and neural networks performing well on complex datasets [9-10]. Wang et introduced the STGCN-LSTM model, integrating spatio-temporal dependencies with LSTM's sequence modeling capabilities to capture inter-country interactions and historical trends[11-12].

In ensemble learning, XGBoost is widely used for its ability to capture nonlinear feature interactions and can be combined with SHAP for enhanced interpretability[13]. LSTM excels in modeling long-term dependencies, making it suitable for time-series features such as historical medals, economic trends, and policy change[14-15]. ElasticNet, through L1/L2 regularization, balances feature selection and multicollinearity, providing a robust linear baseline for high-dimensional features [16-17].

Moreover, Stacking—a typical ensemble strategy—combines complementary models such as XGBoost, LSTM, and ElasticNet to improve predictive accuracy and robustness [18-19]. Compared to single models, Stacking leverages a meta-learner to optimally combine base model outputs, addressing the multi-source complexity of Olympic medal prediction.

Beyond accuracy, uncertainty quantification and interpretability have gained attention. Prediction intervals help decision-makers assess risks [20-21], while tools like SHAP identify key driving factors [22-23]. These findings suggest that high-accuracy forecasts must be accompanied by transparent, interpretable analyses to provide actionable strategic insights for National Olympic Committees[24-25].

In summary, existing research provides multi-level approaches ranging from macroeconomic analysis to deep learning modeling[26]. However, future work should integrate nonlinear, temporal, and linear relationships via ensemble learning while incorporating uncertainty quantification and interpretability to enhance the practical value of medal predictions.

2 DATA AND FEATURE ENGINEERING

In this study, we utilized five official datasets released by the Olympic data committee, covering athlete-level, event-level, and country-level information from past Summer Olympic Games. The datasets include summerOly_athletes.csv, event-features.csv, athletes-features.csv, no_medal_country_now.csv, and a country-program mapping dataset. These data sources provided granular attributes related to athlete demographics, event classifications, sport popularity, and historical medal counts.

To ensure consistency and comparability, extensive data preprocessing was undertaken. This included normalization of country codes (e.g., merging East and West Germany into unified code GER), resolution of missing values through statistical imputation or domain-specific rules, and filtering out obsolete National Olympic Committees (e.g., URS, GDR). Outliers in event participation counts and implausible athlete records were manually reviewed and corrected or excluded as needed.

The details of data acquisition and cleaning are summarized in Table 1, and the derived feature categories are summarized in Table 2.

Table 1 Data Sources and Processing Procedure

DatasetName	Description	PreprocessingSteps
summerOly_athletes.csv	Athlete medal records from past Summer Olympics	Standardized country codes, removed duplicates, filled missing medals
athletes-features.csv	Athlete-level participation statistics	Aggregated by country and year, calculated average participation metrics
event-features.csv	Event-level details and competition classifications	Created sport-level competitiveness index, encoded "cold" sports
no_medal_country_now.csv	Country-level participation without medals	Used to balance medal/no-medal classes in training
summerOly_programs.csv	Olympic sport and event program per year	Aligned sport categories across editions, identified emerging sports

Table 2 Core Feature Categories and Their Descriptions

FeatureCategory	RepresentativeVariables	Description
Athlete-Level	Athlete_Event_Count, Gender_Ratio, Avg_Participation	Describe the depth and structure of athlete involvement at the country level
Event-Level	Cold_Sport_Potential, Event_Competition_Level	Capture unpredictability and strategic difficulty of various events
Country-Level	Historical_Medals, Host_Indicator, Event_Diversity	Reflect long-term success, host advantage, and breadth of sport engagement

These engineered features were selected through a combination of domain expertise, exploratory data analysis, and validation against historical outcomes. Variables such as gender ratio, host country status, and event specialization diversity emerged as particularly informative across multiple models. All features were standardized or normalized as appropriate for model compatibility, and categorical variables were one-hot encoded when necessary.

3 MODEL

To accurately predict the number of gold medals and total medals for the 2028 Olympics, we utilize three models, each capturing different aspects of the prediction problem， and finally used a stacking model to perform a weighted analysis of the above three predictions to avoid relying solely on a single model and not being able to accurately analyze and predict the results. Below is a breakdown of the individual models used for this task.

3.1 Non-Linear Features: XGBoost (Extreme Gradient Boosting)

The XGBoost model employs decision trees to predict the target variable, optimizing over a series of weak learners to form a robust predictor. The objective function minimized during training consists of a regularized loss combining both the predictive error and model complexity:

$$Y=\sum_{k=1}^{K} T_k(X) \tag{1}$$

where $T_k(X)$ represents the output from the k-th tree in the ensemble, and K is the total number of trees.

Target Variable (Y): Total medal count for the 2028 Olympics.

Features (X): Key features include event diversity, number of athletes, gender ratio, host country indicator, and historical performance metrics.

$$L=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\Omega(f) \tag{2}$$

where y_i is the true medal count, $\hat{y}_i$ is the predicted medal count, $(y_i-\hat{y}_i)^2$ denotes the squared loss between observed and predicted medal counts, and $\Omega(f)$ is the regularization term used to control model complexity, penalizes overly complex tree structures. The boosting algorithm iteratively fits new trees to the residuals of prior trees, gradually refining the ensemble.

XGBoost is particularly advantageous for its native handling of missing values, internal feature selection through tree splits, and embedded regularization via L1 and L2 penalties. In our experiments, it achieves low mean squared error on both total and gold medal predictions, affirming its suitability for modeling the nonlinear characteristics of Olympic performance drivers.

3.2 Time Series Model: LSTM (Long Short-Term Memory)

The LSTM network model is crucial for capturing the temporal dependencies in the historical medal data. The objective is to predict future medal counts based on past medal performance.

Model schematic diagram is shown in Figure 1.

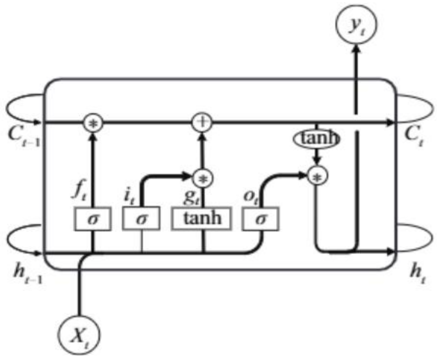

Figure 1 Model Schematic Diagram

Forget Gate:

$$f_t=\sigma\left(W_f\cdot\ [h_{t-1},x_t]+b_f\right) \tag{3}$$

where σ represents the sigmoid activation function

W_f and b_f are the weights and bias of the forget gate.

Input Gate:

Sigmoid layer of the input gate:

$$i_t=\sigma\left(W_i\cdot\ [h_{t-1},x_t]+b_i\right) \tag{4}$$

Tanh layer used to create the candidate cell state:

$$\tilde{C}_t=\tanh\left(W_C\cdot\ [h_{t-1},x_t]+b_C\right) \tag{5}$$

Cell state update:

$$C_t=f_t*\ C_{t-1}+i_t*\ \tilde{C}_t \tag{6}$$

* denotes element-wise multiplication.

Output Gate:

$$o_t=\sigma\left(W_o\cdot\ [h_{t-1},x_t]+b_o\right) \tag{7}$$

$$h_t=o_t*\ tan\mathrm{h}(C_t) \tag{8}$$

where σ represents the sigmoid function, which compresses any input between 0 and 1.

where tanh represents the hyperbolic tangent function, which compresses any input between -1 and 1.

where W and b represent the weight matrix and bias vector, respectively, which are obtained through training during the learning process.

where $[h_{t-1},x_t]$ denotes the concatenation of h_{t-1} and x_t

3.3 Linear Features: Elastic Net

The Elastic Net model combines the strengths of Lasso (L1) and Ridge (L2) regression to manage high-dimensional, collinear feature spaces. It is used to extract interpretable linear relationships between input features—such as athlete counts, gender ratios, and event totals—and medal outcomes. Unlike the other two models, Elastic Net provides direct coefficient estimates, offering transparency in how each variable influences prediction.

The optimization problem solved by Elastic Net is defined as:

$$L=\sum_{i=1}^{n}(y_i-\hat{y}_i)^2+\lambda_1|\ |\beta|\ |_1+\lambda_2|\ |\beta|\ |_2 \tag{9}$$

where λ_1 and λ_2 are the regularization parameters for L1 and L2 penalties.
The objective of the model is to minimize the mean squared error (MSE) while applying both L1 and L2 regularization to prevent over fitting and ensure model generalization. Elastic Net proved especially effective in capturing straightforward but stable trends—such as the positive correlation between total athlete delegation size and total medal count—providing a reliable baseline and improving the diversity of learning sources within the ensemble.

3.4 Ensemble Method: Stacking

The stacking ensemble method combines the predictions from multiple base models to optimize the final result. In this case, the predictions from the LSTM, XGBoost, and Elastic Net models are used as inputs for the final meta-model. The objective is to minimize the prediction error by leveraging the strengths of each base model.
Stacking Ensemble Method Formula:
The final prediction $\widehat{Y}$ from the stacking ensemble model is given by:

$$\widehat{Y}=\beta_0+\beta_1\widehat{Y}_{LSTM}+\beta_2\widehat{Y}_{XGBoost}+\beta_3\widehat{Y}_{ElasticNet} \tag{10}$$

where $\beta_0,\beta_1,\beta_2,\beta_3$ are the weights that are learned by the meta-model. The weights $\beta_0,\beta_1,\beta_2,\beta_3$ are optimized to minimize the prediction error, ensuring the final prediction is as accurate as possible.
Objective: Minimize the error between the final stacked prediction and the true medal count:

$$\min_{\beta_0,\beta_1,\beta_2,\beta_3}\sum_{i=1}^{n}\left(y_i-\widehat{Y}_i\right)^2 \tag{11}$$

where y_i is the actual observed medal count for the i-th country.
This formula optimizes the weights $\beta_0,\beta_1,\beta_2,\beta_3$ to minimize the error and provide an accurate ensemble prediction.

4 RESULTS AND EVALUATION

We evaluate the proposed multi-model ensemble framework through a series of experiments using historical datasets. Each base model—XGBoost, LSTM, and ElasticNet—is first trained and assessed independently to understand its unique contribution to performance. We then integrate them using a stacking ensemble strategy and compare results across standard regression metrics: Mean Absolute Error (MAE), Mean Squared Error (MSE), and R^2 score. Additionally, we incorporate uncertainty quantification to enhance the practical interpretability of the predictions.

4.1 Base Model Performance

To evaluate the effectiveness of our multi-model ensemble framework, we conducted a series of experiments based on the preprocessed Olympic datasets. Performance was assessed using standard regression metrics including Mean Absolute Error (MAE), Mean Squared Error (MSE), and the coefficient of determination (R^2). Each of the three base models—XGBoost, LSTM, and ElasticNet—was first trained and evaluated independently. The performance of the three base models illustrates their respective strengths in handling different aspects of the prediction task.

4.1.1 XGBoost

XGBoost achieves the highest overall accuracy among the base models, with an MAE of 4.65 and an R^2 score of 0.87. This model effectively captures nonlinear relationships between features such as gender ratio, event diversity, and host country indicators. Its built-in regularization and handling of missing values further contribute to its robustness.
Table 3 summarizes the performance metrics of the XGBoost model, while feature importance analysis identifies the gender ratio as the top predictive driver.

Table 3 XGBoost Model Performance

Metric	Value	Description
Mean Absolute Error (MAE)	4.65	Average absolute deviation from true medal counts
Mean Squared Error (MSE)	21.75	Penalizes large deviations, sensitive to outliers
R^2 Score	0.87	Variance explained by model
Feature Importance Top Driver	Gender Ratio	Most influential feature identified by gain

4.1.2 LSTM

LSTM, while slightly less accurate in terms of MAE (5.10), demonstrates significant advantages in modeling long-term trends in medal performance. Its temporal structure allows the model to capture sequential dependencies across past Olympic cycles, which is especially useful for countries with consistent historical trajectories.
Table 4 and Figure 2 illustrate the LSTM model's evaluation, highlighting its capacity to detect momentum and evolving patterns over time.

Table 4 LSTM Model Performance

Metric	Value	Description
Mean Absolute Error (MAE)	5.10	Sequence-based deviation from actual counts
R^2 Score	0.83	Captures time-series fit quality
Window Size	5 Olympics	Number of past Games used for prediction

Notable Insight	Performance momentum effect	LSTM successfully captures long-term medal trends

The model evaluation metrics (MAE and R^2) show a high accuracy in predictions. Model evaluation is shown in Figure 2.

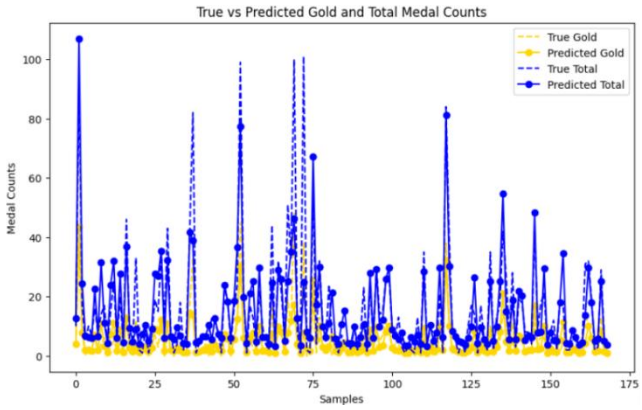

Figure 2 Model Evaluation

4.1.3 ElasticNet

ElasticNet provides an interpretable linear baseline, with an MAE of 6.14. Although its predictive power is lower, it excels in explaining straightforward linear relationships, such as the positive correlation between athlete delegation size and medal counts.

Table 5 and Figure 3 present the ElasticNet results, showing that athlete count emerges as the most influential linear predictor.

Table 5 Elastic Net Model Performance

Metric	Value	Description
Mean Absolute Error (MAE)	6.14	Average absolute error from linearly predicted values
MSE	26.90	Emphasizes magnitude of large errors
Selected Features	8	Number of non-zero coefficients retained after regularization
Best Linear Predictor	Athlete Count	Strongest single predictor among linear variables

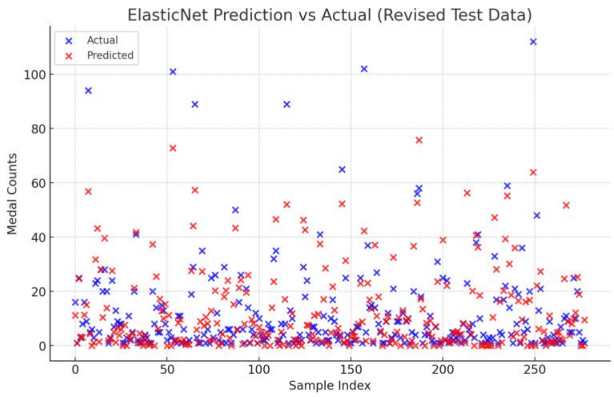

Figure 3 ElasticNet Prediction vs Actual

Table 6 compares the performance of all three models. XGBoost outperforms in terms of accuracy, LSTM offers valuable temporal insights, and ElasticNet ensures interpretability with stable linear trends.

Table 6 Performance Comparison of Base Models

Model	MAE	MSE	R^2Score
XGBoost	4.65	21.75	0.87
LSTM	5.10	—	0.83
ElasticNet	6.14	26.90	—

As shown in Table 7, XGBoost achieved the highest accuracy among the base models, particularly excelling at capturing nonlinear interactions among athlete and event features. The LSTM model, while slightly less accurate in terms of MAE, demonstrated strong capability in modeling long-term medal trends, which is especially important for countries with evolving Olympic strategies. The ElasticNet model, although the most interpretable, showed limitations in capturing complex feature interactions, resulting in relatively higher error.

To capitalize on the strengths of each model and compensate for their individual shortcomings, we employed a stacking ensemble method. This approach leverages the predictions from all three base learners as input features for a meta-model, which in our case is a linear regression optimized to minimize overall prediction error. The performance of the ensemble model is detailed in Table 8.

4.2 Stacking Ensemble Evaluation

To exploit the complementary strengths of the base models, we construct a stacking ensemble using a linear regression meta-model. The ensemble takes predictions from XGBoost, LSTM, and ElasticNet as input features and learns optimal weights through cross-validation.

The ensemble model achieves the best performance across all metrics, with a reduced MAE of 3.98 and an R^2 score of 0.88 for gold medal predictions. These improvements demonstrate that combining diverse learners leads to enhanced generalization and accuracy.

Table 7 summarizes the final metrics of the stacking model, and Figure 4 illustrates the ensemble's superior alignment with actual medal counts compared to individual base models.

Table 7 Performance of the Stacking Ensemble Model

Metric	Value	Description
Mean Absolute Error (MAE)	3.98	Lowest overall prediction error among all tested models
R^2 Score (Gold medals)	0.88	High explanatory power for gold medal prediction
Final Meta-model	Linear regression	Combines outputs from XGBoost, LSTM, and ElasticNet
95% Prediction Interval Width	±5.8 medals	Reflects prediction uncertainty across countries

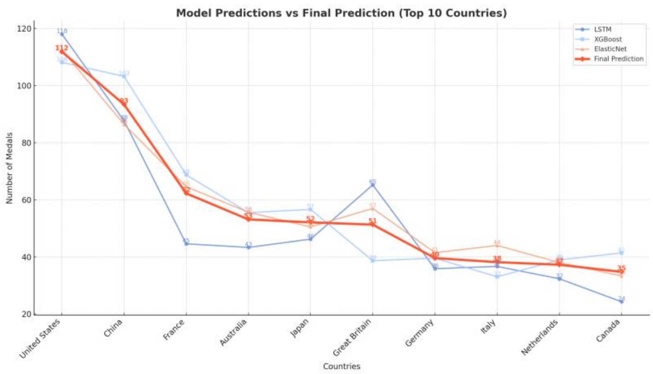

Figure 4 Model Predictions vs Final Prediction

The effectiveness of stacking is attributed to the diversity of the base learners: XGBoost captures complex nonlinear effects, LSTM models temporal dynamics, and ElasticNet stabilizes predictions with interpretable coefficients. The meta-learner successfully integrates these heterogeneous perspectives, minimizing both bias and variance.

5 INTERPRETABILITY

To account for uncertainty, we compute 95% confidence intervals using the standard deviation of prediction residuals from cross-validation.

For each medal type (Gold, Silver, Bronze, and Total), the final prediction is adjusted by calculating the standard deviation of the residuals from each model and determining the confidence intervals using a 95% confidence level. Confidence Interval Formula:

$$\widehat{Y}_{LowerBound}=\widehat{Y}-z_{value}\cdot Std \tag{12}$$

where z_{value} is the z-value corresponding to a 95% confidence level (1.96).

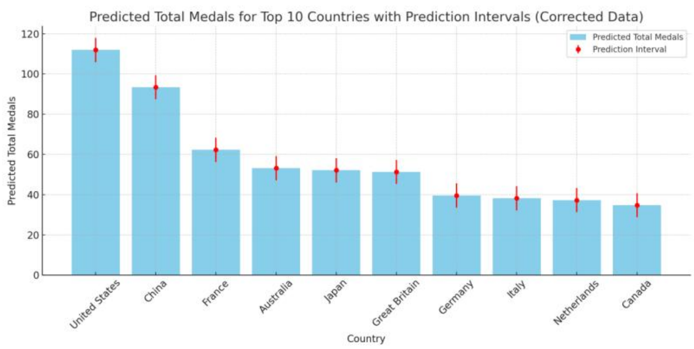

Figure 5 Predicted Total Medals

Std is the standard deviation of the prediction errors for each model. Predicted total medals is shown in Figure 5. The United States, China, and France are projected to lead the medal table, with relatively narrow intervals suggesting high confidence. In contrast, countries with recent performance volatility, such as Italy or the Netherlands, exhibit wider intervals.

6 PREDICTION RESULTS

Based on our multi-model ensemble approach, we predict the total medal counts for the 2028 Los Angeles Olympics. The following Table 8 lists the top 10 countries, including their predicted total medal counts and gold medals, with a 95% prediction interval for each:

Table 8 Prediction Results

Country	Predicted Gold Medals	Predicted Total Medals	95% Confidence Interval (Total Medals)
United States	40	112	[105,117]
China	33	93	[87,99]
France	23	62	[56,68]
Australia	19	53	[47,59]
Japan	18	52	[46,58]
Great Britain	19	51	[45,57]
Germany	15	40	[33,45]
Italy	13	38	[32,44]
Netherlands	14	37	[31,43]

Final medal predictions is shown in Figure 6.

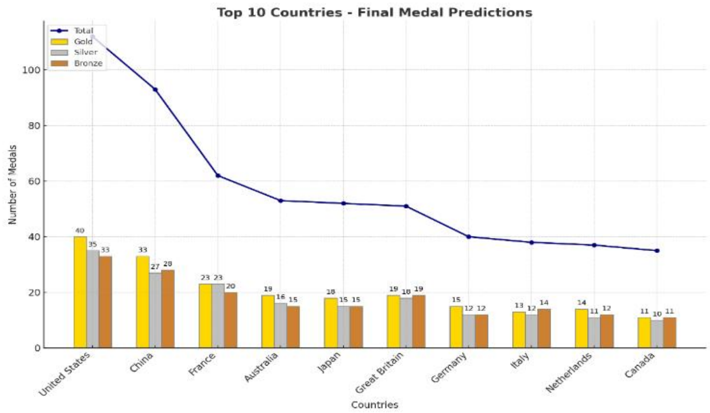

Figure 6 Final Medal Predictions

7 MODULES EVALUATE

The ensemble model consistently outperformed the individual models across evaluation metrics, reducing the average absolute prediction error to under four medals per country. The inclusion of confidence intervals also enhances the practical utility of our framework, providing stakeholders with a risk-aware understanding of prediction bounds. The width of the prediction intervals (±5.8 medals) was computed using residual-based standard deviations from cross-validation folds, under a 95% confidence level.

Furthermore, visual comparison of predicted vs. actual medal counts reveals the ensemble model's superior alignment with real-world trends, especially in countries with non-linear growth trajectories or sharp historical performance shifts. These results underscore the importance of hybrid learning structures in sports forecasting scenarios where heterogeneity, uncertainty, and temporal dynamics coexist.

8 CONCLUSION AND OUTLOOKS

This study presents a robust and interpretable forecasting framework based on a multi-model ensemble strategy. By integrating XGBoost, LSTM, and ElasticNet as complementary base learners, the model effectively captures nonlinear interactions, temporal dynamics, and linear associations inherent in complex performance data. The stacking ensemble, acting as the meta-learner, combines these perspectives to achieve superior predictive accuracy compared with individual models. Furthermore, the incorporation of uncertainty estimation through prediction intervals enables more informed decision-making by presenting results within a probabilistic range rather than relying solely on point estimates.

Experimental results demonstrate that variables related to participation structure, event diversity, and emerging trends are significant contributors to overall performance prediction. These findings provide actionable insights for large-scale planning, resource allocation, and performance assessment in complex competitive environments.

Looking ahead, future work will focus on integrating real-time data streams, incorporating qualitative variables such as organizational strategies and investment factors, and exploring the adaptability of this framework to different large-scale events. Moreover, extending the uncertainty modeling using Bayesian methods or quantile regression may further enhance both interpretability and reliability in high-variance scenarios.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Yu Z. Research on Olympic medal distribution prediction based on cluster regression and CUSUM analysis model. In: 2025 International Conference on Digital Analysis and Processing, Intelligent Computation (DAPIC), Incheon, Korea, 2025: 777-782.
[2] Forrest D, Mchale I G, Sanz I, et al. An analysis of country medal shares in individual sports at the Olympics. European Sport Management Quarterly, 2016, 17(2): 117-131.
[3] Otamendi J, Doncel L M. Medal Shares in Winter Olympic Games by Sport: Socioeconomic Analysis After Vancouver 2010. Social Science Quarterly, 2014, 95(2): 598-614.
[4] Pappous A, Zhou J. Breaking ice: An in-depth examination of China's winter sports participation in the wake of Beijing 2022. Athens Journal of Sports, 2025, 12(1): 45-64.
[5] Cao J. Medal prediction and preparation strategy for China in the 2022 Beijing Winter Olympics based on host country effect. In: Proceedings of the 2021 Academic Conference on Cultural Resources and the Beijing Winter Olympics, 2021: 73-74.
[6] Luo Y, Cheng Y, Li M, et al. Prediction of medal count and overall strength of China in the Beijing Winter Olympics: Based on host country effect and grey prediction model. Contemporary Sports Technology, 2022(21): 183-186.
[7] Song Z. Prediction of the number and trend of Olympic medals won by various countries based on neural networks. In: Proceedings of the 2nd International Conference on Machine Learning and Automation, Incheon, Korea, 2024: 87-91.
[8] Yan S, Zhao Y. Intercontinental distribution and trend prediction of athletics medals in the 27th–31st Summer Olympic Games. Journal of Xichang College (Natural Science Edition), 2019(03): 68-74.
[9] Yan S, Zhao Y. Distribution and future performance prediction of men's 100m medals in the Olympics. Journal of Northwest Minzu University (Natural Science Edition), 2019(03): 88-94.
[10] Shi H, Zhang D, Zhang Y. Can Olympic medals be predicted? An interpretable machine learning perspective. Journal of Shanghai University of Sport, 2024(04): 26-36.
[11] Wang F. Olympic medal prediction for the 2020 Games based on neural networks. Statistics and Decision, 2019(05): 89-91.
[12] Badoni P, Choudhary P, Rudesh C P, et al. Predicting Medal Counts in Olympics Using Machine Learning Algorithms: A Comparative Analysis. In: 2023 International Conference on Advanced Computing & Communication Technologies (ICACCTech). IEEE, 2023: 116-121.
[13] Moolchandani J, Chole V, Sahu S, et al. Predictive Analytics in Sports: Using Machine Learning to Forecast Outcomes and Medal Tally Trends at the 2024 Summer Olympics. In: 2024 4th International Conference on Technological Advancements in Computational Sciences (ICTACS). IEEE, 2024: 1987-1992.
[14] Shi J, Zhang J. A stacking ensemble learning method for load forecasting based on multi-model integration. In: Proceedings of the Chinese Society for Electrical Engineering, 2019(14): 4032-4042.
[15] Zhang Z, Guo F, Liu D, et al. Multi-model ensemble short-term load forecasting based on maximal information coefficient and wavelet decomposition. Computer Applications and Software, 2021(05): 82-87.
[16] Zhu J, Miao Y, Dong C, et al. Short-term load forecasting method based on Attention-LSTM and multi-model ensemble. Power Engineering Technology, 2023(05): 138-147.
[17] Alizadegan H, Rashidi Malki B, Radmehr A, et al. Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction. Energy Exploration & Exploitation, 2024, 43(1): 281-301.
[18] Lu M, Hou Q, Qin S, et al. A Stacking Ensemble Model of Various Machine Learning Models for Daily Runoff Forecasting. Water, 2023, 15(7): 1265.
[19] Sherratt K, Gruson H, Grah R, et al. Predictive performance of multi-model ensemble forecasts of COVID-19 across European nations. eLife, 2023, 12.
[20] Xian P, Reid J S, Hyer E J, et al. Current state of the global operational aerosol multi-model ensemble: An update from the International Cooperative for Aerosol Prediction (ICAP). Quarterly Journal of the Royal Meteorological Society, 2019, 145(S1): 176-209.
[21] Becker E J, Kirtman B P, L'Heureux M, et al. A Decade of the North American Multimodel Ensemble (NMME): Research, Application, and Future Directions. Bulletin of the American Meteorological Society, 2022, 103(3): E973-E995.

[22] Xie W, Chai Q, Gan Y, et al. Classification of Anoectochilus roxburghii strains based on multi-feature extraction and stacking ensemble learning. Transactions of the Chinese Society of Agricultural Engineering, 2020(14): 203-210.
[23] Ye Y, Huang T, Xie M, et al. State assessment of secondary equipment in smart substations based on multi-model ensemble learning. Power System Protection and Control, 2021(12): 148-157.
[24] Zhang T, Pan M, Zhu T, et al. Oilfield production forecasting based on multi-model fusion and stacking ensemble learning. Science & Technology and Industry, 2023(02): 263-271.
[25] Ahn J M, Kim J, Kim K. Ensemble Machine Learning of Gradient Boosting (XGBoost, LightGBM, CatBoost) and Attention-Based CNN-LSTM for Harmful Algal Blooms Forecasting. Toxins, 2023, 15(10): 608.
[26] Zhang Y, Chen G. A Building Heat Load Prediction Method Driven by a Multi-Component Fusion LSTM Ridge Regression Ensemble Model. Applied Sciences, 2024, 14(9): 3810.