

Beyond Medals: A Predictive Model for Olympic Success and Identifying Opportunities

Summary

The Olympics have always been a topic of great interest. With the 2028 Summer Olympics set to take place in Los Angeles, there is much anticipation not only for the exciting competitions but also for which countries will achieve success or experience setbacks in the medal standings. Meanwhile, a greater and more intriguing challenge has emerged: utilizing information from past Olympics to predict future Olympic medal outcomes. This is the primary question we address in this paper.

To establish a reasonable model, our first task is to develop an appropriate model for national medal counts. We built a bottom-up medal analysis and prediction model using the **TrueSkill-SARIMAX-RF method**. For traditional events, the **TrueSkill** model is used to evaluate a country's performance in each event. Then, the **SARIMAX** model is applied to predict the country's performance level for 2028. The host country factor was considered in this step. To account for potential errors, a Bayesian prediction method is employed. Finally, the predicted performance for 2028 determines the medal allocation. The results from these two parts of events are aggregated to form the final medal count model for each country.

For Task 1, using our model, we predicted medal winners for each event. Based on our simulated medal table, we also predicted which countries are most likely to show the greatest decline or improvement. Our model then address the second major question in Task 1: predicting medal-winning probabilities for countries that have never won medals. Through rigorous probability analysis, we concluded that in 2028, there is a 0.25 probability that 2 countries without previous medals will win their first medals. Regarding the third question in Task 1, since we derived overall national medal outcomes from event-level predictions, we aggregate event performance predictions to determine which sports are most crucial for each country. Finally, we analyzed the host country's newly chosen events and various countries' medal performances, discovering that host countries have higher chances of winning medals in home competitions and newly selected events.

For Task 2, we performed **residual analysis** to find evidence of how coaches influence athletes' performance and medal counts. By analyzing different variables, we identified anomaly information in the data, providing evidence of how the possible "great coach" effect alters the data. Subsequently, we introduced **LR-SCAD** model to study how the "great coach" effect contributed to changes in the total medal count for different countries and created visualizations for this. Finally, we selected **USA NOC**, **SWE NOC**, and **CHN NOC** and identified the sports in which these countries should invest in a great coach. We then evaluated the impact of the great coach on their medal counts.

For Task 3, based on the model, we analyzed and derived 2 factors that affect a country's medal count, which were not covered in Tasks 1 and 2. We found that countries without previous medals should introduce coaches in events with higher winning probabilities to significantly increase their chances of success. Additionally, we discovered that the proportion of experienced athletes in a team affects medal-winning probabilities.

Lastly, we conducted **sensitivity analysis** on our model, finding that it demonstrates strong generalizability across different datasets, proving its robustness.

Keywords: TrueSkill, SARIMAX, Residual Analysis, LR-SCAD Model, Great Coach Effect

Contents

1	Introduction	3
1.1	Problem Background	3
1.2	Restatement of The Problem	3
1.3	Our Work	4
2	Assumptions	5
3	Notations	5
4	Data Preprocessing	6
5	Medal Projection Model	6
5.1	TrueSkill Sports Performance Evaluation Model	7
5.2	SARIMAX-Bayesian Model: Predicting Future Medals	8
5.3	Random Forest Model: Predicting Medals in New Events	10
5.4	Model Aggregation	11
5.5	Model Evaluation	11
6	Task 1: Medal Prediction and Analysis for the 2028 Olympics	12
6.1	Medal Table Projections	12
6.2	First Ever Medal	13
6.3	Task 1.3	14
7	Task 2: Great Coach Effect	16
7.1	Residuals Test: Evidence of Changes That Might Be Due to Great Coach Effect	16
7.2	LR-SCAD: Estimate Great Coach Effect Contributes To Medal Counts	18
7.3	Choose three countries and identify sports where they should consider investing in a “great” coach and estimate that impact	19
8	TASK 3: Oringinal Insights	20
8.1	Insights and Recommendations for Countries That Have Never Won Medals	20
8.2	Insights and Advice for NOCs on Deploying Experienced Athletes	21
9	Model Evaluation	22
10	Sensitivity Analysis	23
11	Conclusion	24
	References	25



1 Introduction

1.1 Problem Background

The Olympics have always been a topic of great interest. With the 2028 Summer Olympics set to take place in Los Angeles, there is much anticipation not only for the exciting competitions but also for which countries will achieve success or experience setbacks in the medal standings, as well as whether countries that have never won medals will secure their first one. Also, predicting the medal counts for the United States in the 2028 Olympics, as the host nation, is an exciting and dynamic challenge.

Figure 1 shows new events added to the 2028 Los Angeles Olympics:



Figure 1: New events added to the Los Angeles Olympics[1].

1.2 Restatement of The Problem

We have divided the entire problem into three main tasks:

Task1: Medal Prediction and Analysis for the 2028 Olympics For Task 1, the following objectives must be accomplished: Build a model to represent and predict the medal counts for each country, including gold medals and total medals, while evaluating the model's accuracy and performance. Once the model is constructed, the subsequent tasks based on its outputs must be completed:

- Predict the medal table for the 2028 Los Angeles Olympics and obtain prediction intervals. Identify the countries most likely to improve or decline in the medal table according to our predictions.
- Predict how many countries that have never won a medal in the Olympics will win medals in the 2028 Los Angeles Olympics, along with the probability of this occurrence.
- Identify the most important sports for different countries, provide reasons, and analyze how the events chosen by the host country influence the gold and total medal counts of various countries.

Task 2 The Impact of the “Great Coach” Effect For Task 2, the study of the “great coach” effect is required. First, evidence needs to be found showing how the “great coach” effect changes the data.

Then, analyze the contribution of the “great coach” effect to changes in the total medal count. Lastly, select three countries, identify which sports they should invest in a “great” coach, and assess the impact on their medal counts.

Task 3 Exploring Strategic Insights For Task 3, additional insights need to be derived from the model, and these insights must be explained in a way that they can provide useful reference for the national Olympic committees.

1.3 Our Work

Figure 2 below shows the implementation steps of our entire work:

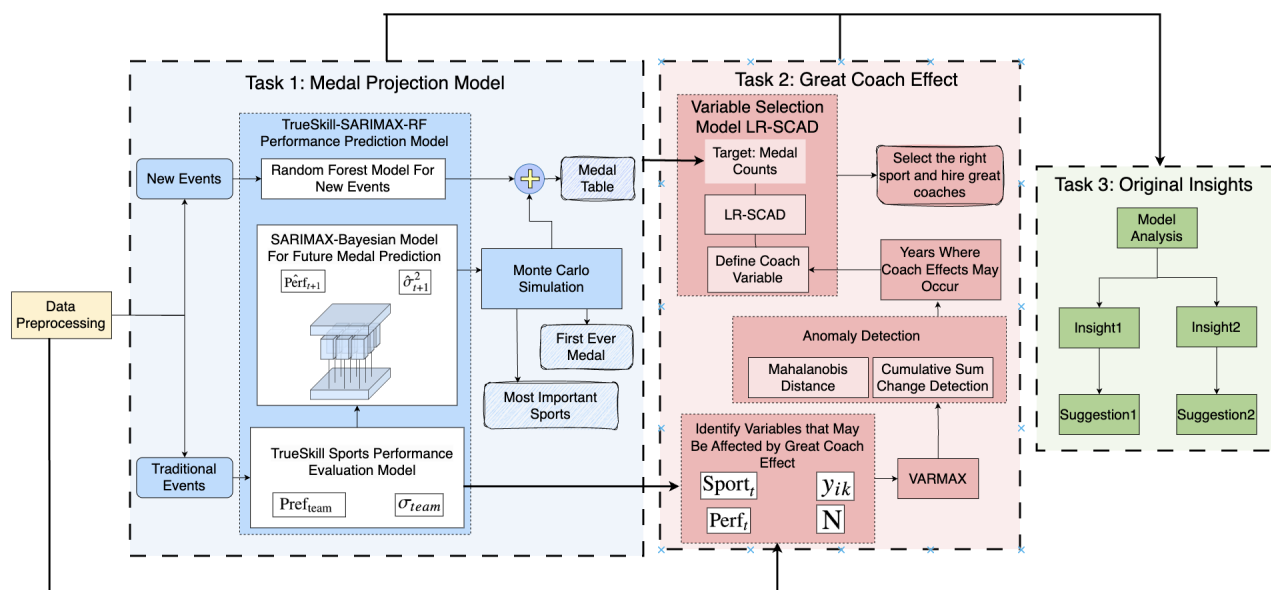


Figure 2: Overall Structure of Our Model

For **Task 1**, we built a bottom-up medal projection model using the TrueSkill-SARIMAX-RF method. The model first categorizes all events into traditional events and new events. For traditional events, the TrueSkill-SARIMAX model is applied to predict the country’s performance level for 2028. To account for potential errors, a Bayesian prediction method is employed. The model determines the allocation of medals based on the predicted performance rankings. For new events, a Random Forest model is used to predict medal allocation. The results from these two parts are aggregated to form the final medal count model for each country. We used the model to predict and analyze all the issues mentioned in the statement of the problem, and provided an analysis of the model’s accuracy.

For **Task 2**, we performed residual analysis. By analyzing different variables, we identified outliers in the data, providing evidence of how the “great coach” effect alters the data. Subsequently, we studied how the “great coach” effect contributed to changes in the total medal count for different countries and created visualizations for this. Finally, we selected USA, CHN and SWE to identify the sports in which these countries should invest in a “great” coach. We then evaluated the impact of the “great” coach on their medal counts.



For **Task 3**, Based on the model, we analyzed and identified several factors that affect a country's medal count, which were not covered in Tasks 1 and 2. These factors were then integrated into insights, and based on these insights, we provided recommendations for the concerned national Olympic committees.

2 Assumptions

To ensure the accuracy of our proposed model, we make the below assumptions, which are properly justified:

1. The performance in different sports does not affect each other. This assumption aligns with the International Olympic Committee's sport categorization framework that treats disciplines as distinct competitive domains.
2. All athletes without medals can be treated as holding equivalent competitive standing to fourth-place finishers, because of limited historical records.
3. Each country is restricted to obtaining at most one medal per event, regardless of the number of participating athletes, which enables holistic evaluation of national athletic ecosystems by preventing single-event dominance from overshadowing broader capability distributions.
4. Since dissolved states without internationally recognized successor entities can contaminate modern Olympics prediction result, they are excluded from the modeling framework, with their historical data truncated at the dissolution timestamp.
5. Consistent with IOC sanctions on Russia, Russian athlete participation parameters for the 2028 Olympiad can be excluded, reflecting suspension measures in response to geopolitical conflicts.

3 Notations

The key mathematical notations used in this paper are listed in **Table 1**:

Table 1: Notations used in this paper

Symbol	Description	Symbol	Description
$\Delta\mu$	Difference in Athlete Performance	σ_{team}	National Team Std Dev
$Perf_i^t$	Athlete's Real Performance	ϵ_t	White Noise Error Term
$Perf_{team}$	Team Performance	$Perf_i^{Sports_k}$	Country Sport Performance
$\sigma_{athlete}$	Athlete Std Dev	$Coach_{i,k}$	Coaching Effect

4 Data Preprocessing

Our data preprocessing work primarily focused on the **summerOly_athletes.csv** dataset.

1. We removed data from countries that no longer exist and those no longer competing, like RUS. For merged countries like East and West Germany, we merged them under a single NOC.
2. To better utilize the data, we standardized event names. Events were categorized into currently existing and discontinued events. Discontinued events were removed from the dataset. For existing events, we mapped historical names to their 2024 counterparts.
3. Based on the assumptions outlined in section 2, we consolidated athletes from different teams of the same country into their respective national teams.

5 Medal Projection Model

Figure 3 below shows the implementation steps of Task 1:

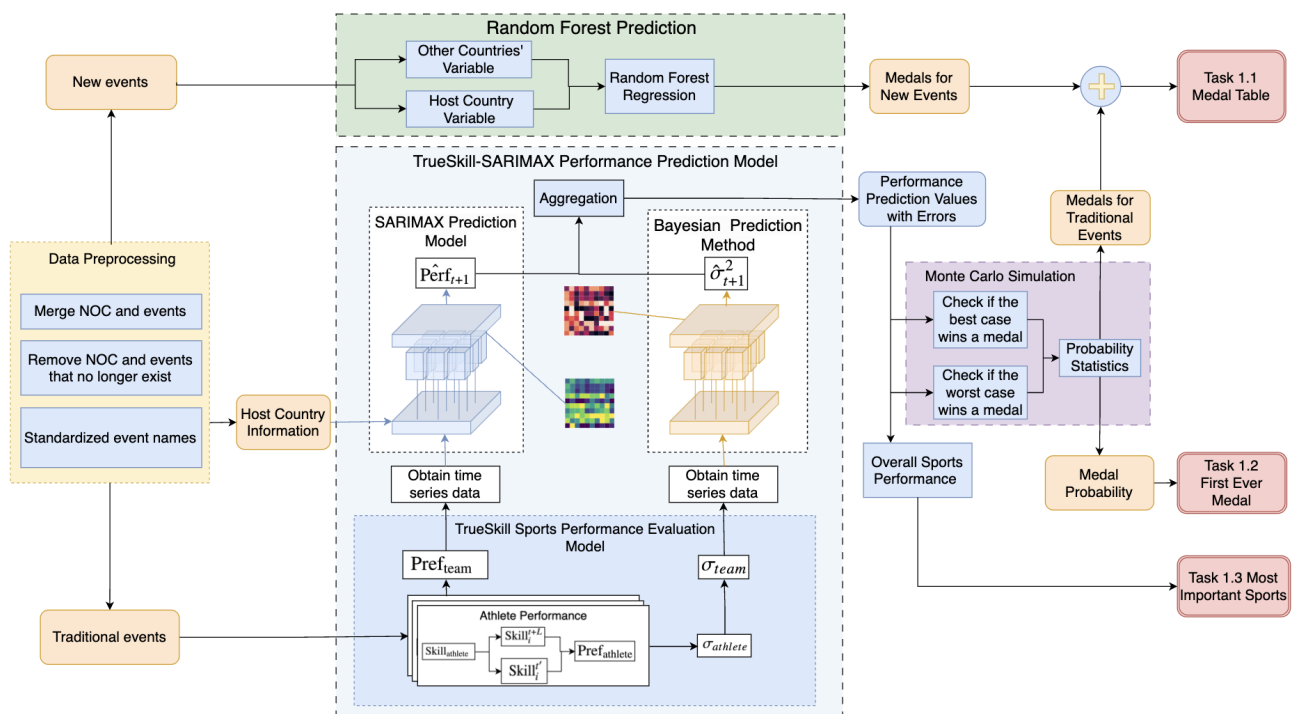
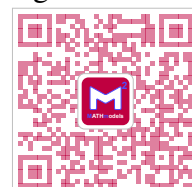


Figure 3: Overall Structure of Task 1

We first divide all events into two categories: traditional events that occur in every Olympics and new events being introduced every year.

For traditional events, we use a TrueSkill-SARIMAX model with a bottom-up approach. This involves evaluating each national team's performance at the event level, predicting their future performance, and determining medal outcomes based on predicted performance rankings.



For new events, we use a random forest model to predict medal outcomes for 2028's new events, using historical data from previously introduced events and their medal distributions.

Finally, we aggregate the medal predictions and their confidence intervals from both categories to produce the overall medal table.

5.1 TrueSkill Sports Performance Evaluation Model

Research indicates that Olympic medal predictions cannot rely solely on historical medal tables. A country's success in specific events depends on their relative **Performance** level. To evaluate a country's performance and uncertainty in various events, we innovatively introduced the **TrueSkill**[2] sports performance evaluation model.

For each event, athletes are grouped by their national teams. Assume each athlete i has a real-valued skill at time t , denoted as skill_i^t . The initial Skill is drawn from a Gaussian distribution with mean m_0 and variance v_0 , expressed as:

$$\text{Skill}_i^{t_0} \sim \mathcal{N}(m_0, v_0) \quad (1)$$

where t_0 represents the athlete's first Olympic appearance time, and m_0 and v_0 are adjustable parameters.

After each Olympics, each athlete's Skill changes by a value drawn from a Gaussian distribution. Athletes also gain experience from competitions:

$$\text{Skill}_i^{t+L} \sim \mathcal{N}(\text{Skill}_i^t, \gamma^2) \quad \text{Skill}_i^{t'} \sim \mathcal{N}(\text{Skill}_i^t, \tau^2(t' - t)), \quad \text{where } t' > t \quad (2)$$

where L represents the number of Olympics participated.

The next step generates match results based on athlete skills. TrueSkill assumes each athlete has a real-valued performance in each match:

$$\text{Perf}_i^t \sim \mathcal{N}(\text{skill}_i^t + \text{teamOffset}(\text{size of team}), \beta^2) \quad (3)$$

where β is an adjustable parameter reflecting game randomness.

For subsequent discussion, we omit Perf_i 's dependence on t as we're discussing specific matches. During Olympic competitions, each team's performance is defined as the weighted sum of all athletes' performances, with weights proportional to their participation time to account for experience:

$$\text{Perf}_{\text{team}} = \sum_{i \in \text{team}} \frac{\text{Perf}_i \cdot \text{timePlayed}_i}{L} \quad (4)$$

Then rank the national performance $\text{Perf}_{\text{team}}$ in descending order to determine which countries can win medals. We adjust m_0 and v_0 according to the actual results of that year. Based on this process, we obtained each country's performance in a particular event for each year.

We also need to determine the error of the event performance. For the error calculation of individual athlete's Skill_i , it is divided into two cases: case 1 is applied when comparing different athletes, while case 2 is applied due to long-term non-participation:

$$\sigma_{\text{athlete}} = \begin{cases} \sqrt{\sigma^2(1 - \frac{\sigma^2}{v_{\text{total}}} \cdot w)}, & w = \frac{\phi(x)}{\Phi(x)} \left(\frac{\phi(x)}{\Phi(x)} + \frac{\Delta\mu}{\sqrt{v_{\text{total}}}} \right), \text{ case 1} \\ \sqrt{\sigma^2 + (\tau \cdot t)^2}, & \text{case 2} \end{cases} \quad (5)$$

where v_{total} represents the total variance, calculated as $v_{\text{total}} = \sum_{i=1}^n \sigma_i^2 + 2\beta^2$, τ is the time decay factor, t is the number of days without participation in the competition.

Subsequently, the national team's standard deviation is derived from the standard deviation values of all athletes within the team:

$$\sigma_{team} = \sqrt{\sum_{i=1}^n \sigma_{player,i}^2} \quad (6)$$

In this way, we obtain the yearly **Performance** values $Perf_t$ for each country in each event and its standard deviation σ_{team} .

Part of our results is illustrated in the following **Figure 4**, showing changes over time in the performance of different countries in various events.

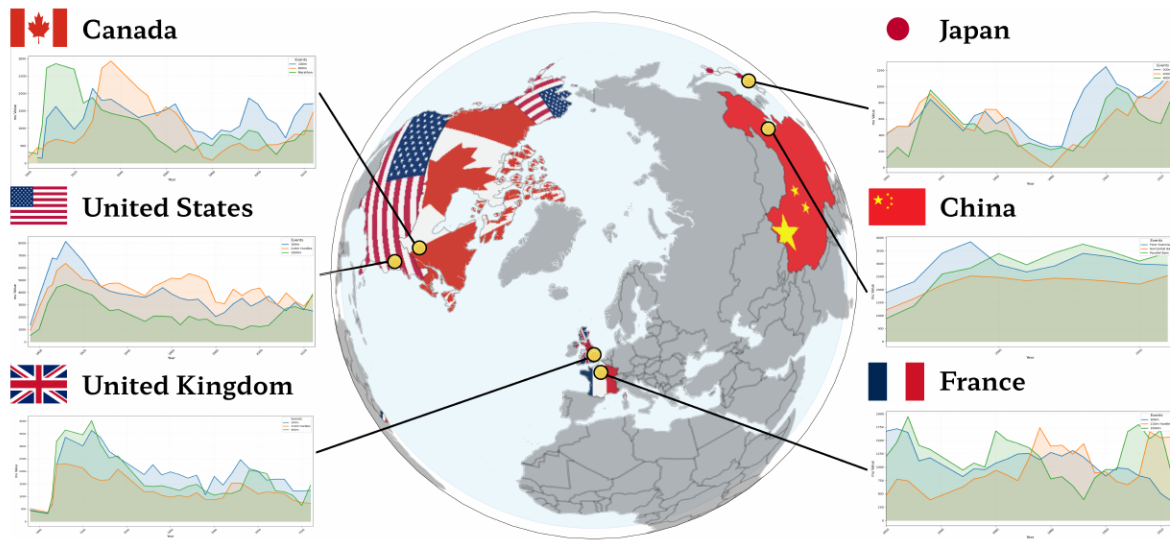


Figure 4: Performance of Different Countries in Various Events

5.2 SARIMAX-Bayesian Model: Predicting Future Medals

From subsection 5.1, we obtained the performance levels of different countries in each Olympic event from 1896-2024. Since a country's performance level in an event shows periodic fluctuations and is directly related to the Olympic host country, we innovatively introduced the SARIMAX time series prediction model to predict future national performance levels in specific events using these factors.

Figure 5 below shows the implementation steps of our SARIMAX-Bayesian Prediction Model:

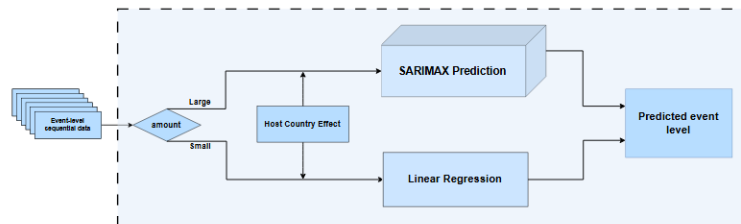


Figure 5: Structure of SARIMAX-Bayesian Prediction Model



关注数学模型
获取更多资讯

The SARIMAX model combines autoregression **AR**, moving average **MA**, seasonal patterns **S**, and exogenous variables **X**[3]. It requires target, a country's performance on a typical event, donated as Perf_t , where t represents time. Host country, donated as Host_t , is the exogenous variable model use.

To begin with, SARIMAX process the AR part, representing the relationship between the current target value and its past values. Also, MA part is processed to represent the relationship between current error and past errors. Periodic patterns are then captured by SARIMAX through seasonal auto-regression coefficients ϕ_k :

$$\text{Perf}_t = \alpha + \sum_{i=1}^p \beta_i y_{t-i} \quad \text{Perf}_t = \epsilon_t + \sum_{j=1}^q \theta_j \epsilon_{t-j} \quad \text{Perf}_t = \sum_{k=1}^S \phi_k y_{t-k \cdot s} \quad (7)$$

Where: α is the constant term. β_1, \dots, β_p are autoregression coefficients, representing the relationship between the current value and previous p moments values. $\theta_1, \dots, \theta_q$ are moving average coefficients that represent the relationship between current error and previous q moment errors. s is the period of time, and ϕ_k is the seasonal autoregression coefficient.

Then, fit the AR and MA coefficients by minimizing the sum of squared residuals:

$$\hat{\beta}_i, \hat{\theta}_j, \hat{\phi}_k, \hat{\gamma} = \arg \min \sum_{t=1}^T (\text{Perf}_t - \hat{\text{Perf}}_t)^2 \quad (8)$$

This completes the model training. Finally, aggregate all parts to obtain the complete model and use the complete model to predict future data:

$$\hat{\text{Perf}}_{t+h} = \alpha + \sum_{i=1}^p \hat{\beta}_i \hat{y}_{t+h-i} + \sum_{j=1}^q \hat{\theta}_j \hat{\epsilon}_{t+h-j} + \sum_{k=1}^S \hat{\phi}_k \hat{\text{Perf}}_{t+h-k \cdot s} + \text{Host}_{t+h} \hat{\gamma} \quad (9)$$

Additionally, the SARIMAX model performs very inaccurately on small datasets. For events with shorter time spans, we used simple linear regression.

Through this method, we obtain the performance levels of events of each country $\hat{\text{Perf}}_{t+1}$ for traditional events in 2028.

In section 5.1, the result $\hat{\text{Perf}}_{t+1}$ of the TrueSkill model involves a standard deviation. Similarly, the prediction results from SARIMAX also have a standard deviation. Since this standard deviation cannot be predicted using the SARIMAX model, we use a Bayesian prediction method, based on the same principle as TrueSkill, to predict the standard deviation.

We model the future standard deviation using an Inverse Gamma distribution, which is a natural conjugate prior for variance parameters[4]:

$$\sigma^2 \sim \text{InvGamma}(\alpha, \beta) \quad (10)$$

Given historical standard deviations $\sigma_1^2, \sigma_2^2, \dots, \sigma_T^2$, we can update the parameters as:

$$\alpha_T = \alpha_0 + \frac{T}{2}, \quad \beta_T = \beta_0 + \frac{1}{2} \sum_{t=1}^T (\sigma_t^2 - \mu)^2 \quad (11)$$

Finally updating α_T and β_T to obtain the new predicted standard deviation:

$$\hat{\sigma}_{T+1}^2 = \frac{\beta_T}{\alpha_T - 1} \quad (12)$$

Using this method, we can obtain the predicted values for the standard deviation $\hat{\sigma}_{t+1}$ of the performance level $\hat{\text{Perf}}_{t+1}$. We made a heatmap to show the predicted performance of certain countries in various events, as shown in the following **Figure 6**. We also made a heatmap to illustrate the predicted standard deviation of certain countries in various events, as shown in the following **Figure 7**:

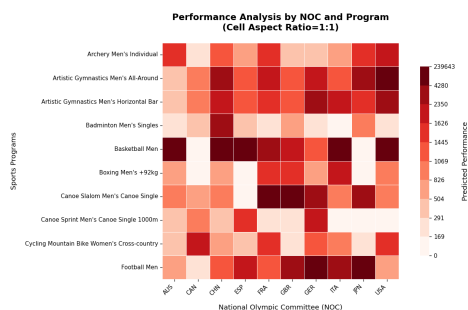


Figure 6: Predicted Performance of Certain Countries in Various Events

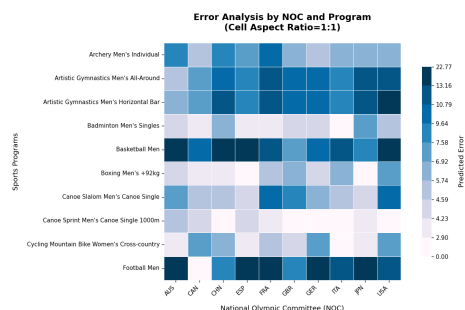


Figure 7: Predicted Standard Deviation of Certain Countries in Various Events

5.3 Random Forest Model: Predicting Medals in New Events

For newly added events, since the new events in each Olympic Games are uncertain, and independent from other events, they cannot be evaluated using the TrueSkill model or predicted with SARIMAX. Therefore, we applied the Random Forest model to predict the number of medals in new events.

Random Forest regression[5] is an ensemble learning method that builds multiple decision trees for regression prediction. Each tree is trained on a different subset of the data and features, and the final prediction is obtained by averaging the results from all trees.

Our model should first determine the input variables. Input variables include the total number of medals in both new and traditional events for each Olympic Games, the number of medals won by the host country in these two types of events, and the number of medals won by other countries in each of these events.

These variables are then fed into the Random Forest model, which outputs predictions and model errors on test cases, including MSE and MAE. We selected the best-performing parameters to complete the model training. The errors on the test set are shown in **Table 2**:

Table 2: Performance Metrics of Random Forest Models	
Target Variable	MAE
Host Gold	0.1988
Host Silver	0.3460
Host Bronze	0.2784

It can be observed that our model does have a certain degree of accuracy.

Finally, we used the medal predictions for traditional events in 2028 (from sections 5.1 and 5.2) as inputs to predict the performance in new events for each country.



关注数学模型
获取更多资讯

5.4 Model Aggregation

The predictions from the 3 sections above are then aggregated to predict medals. By adding the medal counts of traditional events and newly added events, we obtain the predicted medal values. However, we must also determine the prediction intervals for the medals. The method is as follows:

For a specific event, for the predicted Pref from Section 5.1, we have $\text{Pref}_t \sim \mathcal{N}(\mu, \sigma^2)$. Therefore, the actual Pref_t of the country will always lie within the interval $(\text{Pref} - 3\sigma, \text{Pref} + 3\sigma)$. From this, we can obtain the best possible performance and the worst possible performance, denoted as Pref_{\min} and Pref_{\max} .

For a given country, the worst-case scenario occurs when its performance is Pref_{\min} and the performance of other countries is Pref_{\max} . We use these $\text{Pref}_{\text{extreme}}$ to reorder the countries in descending order and determine whether the country will win a medal in that event. This process is repeated for all countries in all events to obtain a worst-case medal table. Similarly, we can use this method to obtain the best-case medal table.

By combining the two medal table above, we obtained our predicted medal table with prediction intervals.

5.5 Model Evaluation

We use the 2024 data as a test set to evaluate the accuracy of our model. We applied our model to data from the 2020 Olympics and earlier, obtaining medal count prediction intervals. Then, we compared these intervals with the actual values from 2024, resulting in the following Figure:

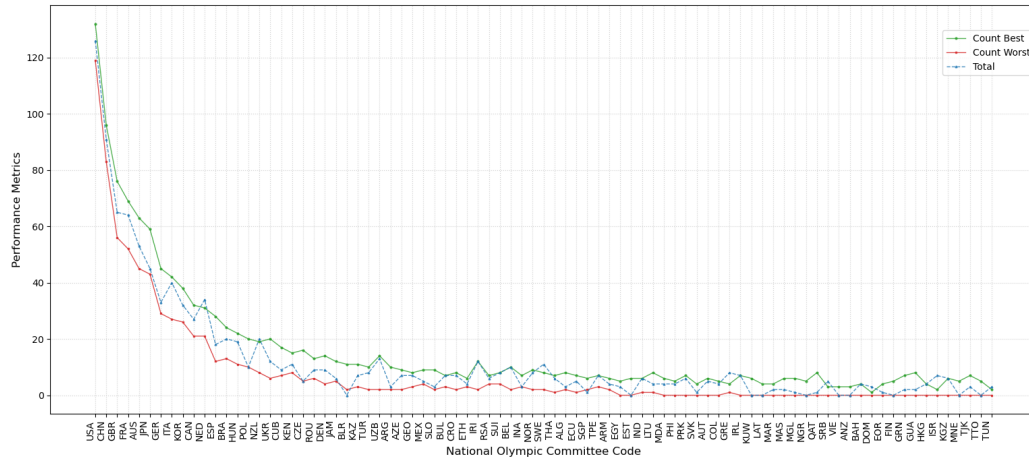


Figure 8: Compare Prediction Intervals With The Actual Values From 2024

As shown in the figure above, our model predictions generally fall within the correct range, demonstrating the accuracy of our model.

6 Task 1: Medal Prediction and Analysis for the 2028 Olympics

We have completed the model construction in Section 5. In the following section, we are going to focus on solving the task.

6.1 Medal Table Projections

Using data from 2024 and earlier Olympics as input to the model, we obtained the virtual medal table with intervals, as shown in **Figure 9** below:

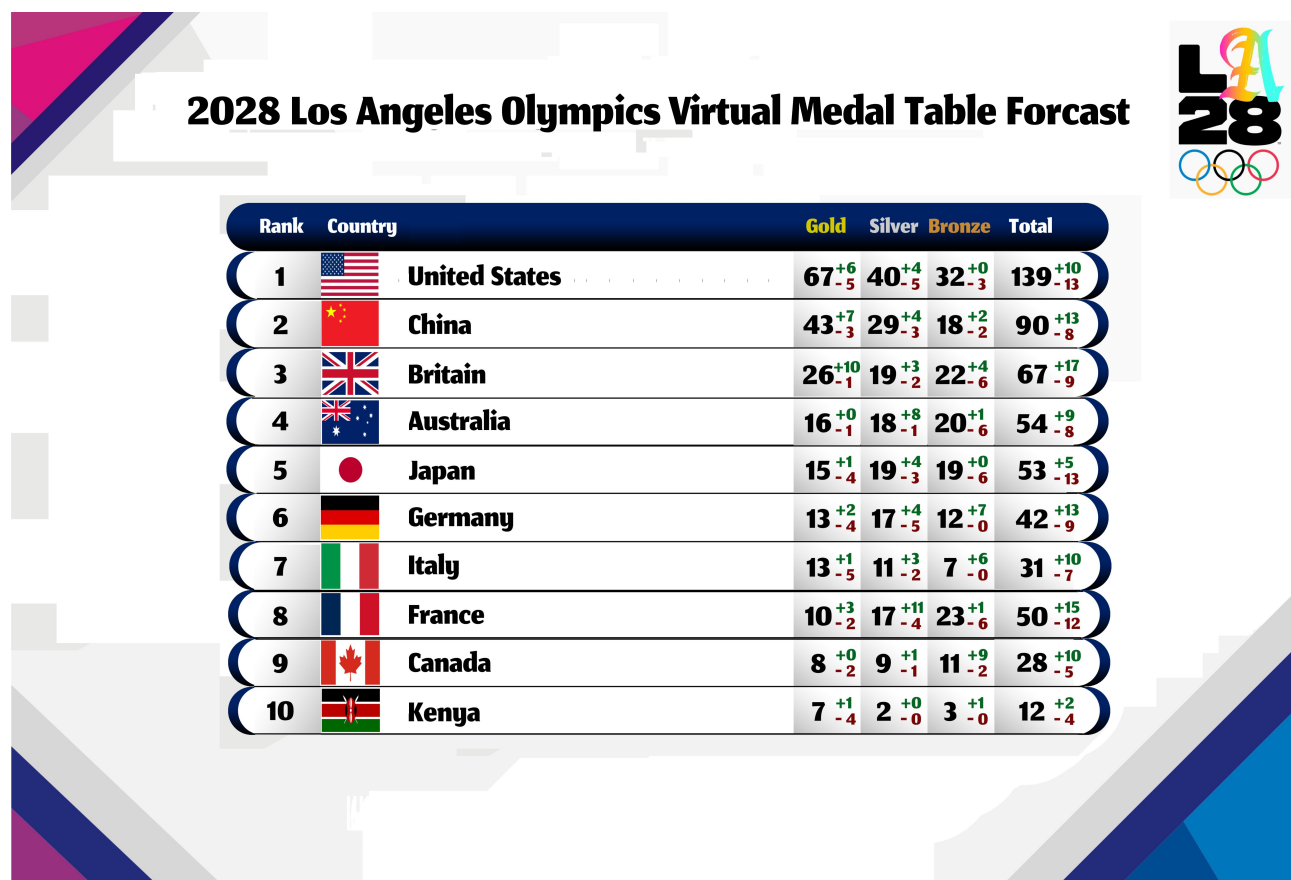


Figure 9: 2028 Virtual Medal Table

We then compare the medal table we obtained with the 2024 medal table and identify the countries that are most likely to improve or worsen in the medal table and gold medal table, as shown in **Figure 10** and **Figure 11** below:



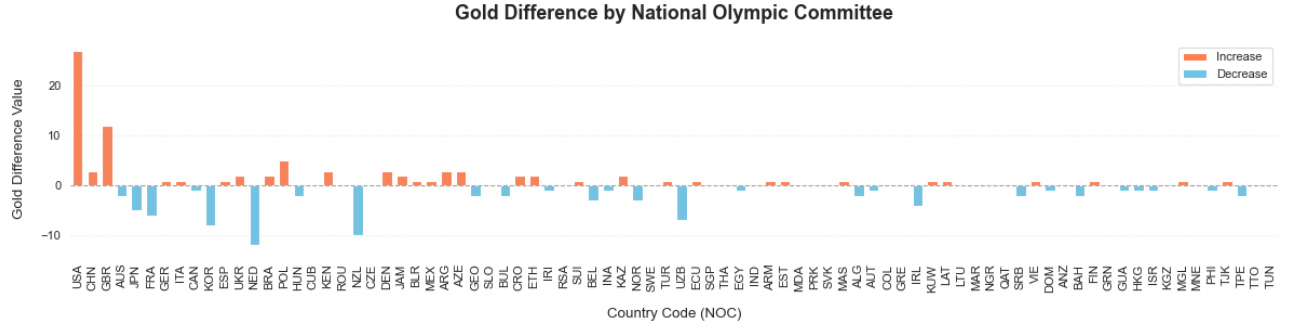


Figure 10: Gold Difference by NOC

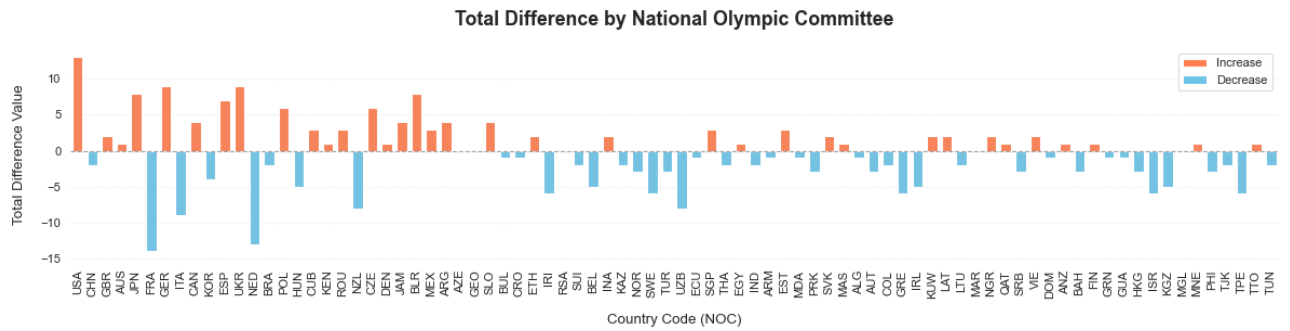


Figure 11: Total Difference by NOC

It is evident that the USA, GBR, and POL are likely to see significant improvements in their gold medal counts, while FRA and NED are likely to experience a decline in their total medal counts.

6.2 First Ever Medal

We have just focused on the top positions in the predicted medal table. Now, let us shift our attention to the bottom of the predicted medal table. By analyzing the historical data, we find that the countries AND and SOL, which had never earned medals in the past, earned medals in our predictions. They each won their country's first medal in the events Athletics Men's 3000m Steeplechase and Athletics Women's 100m, respectively. This suggests that our model can predict medal results for countries that have never won medals before.

So, how many countries are most likely to win medals in 2028? We conducted a Monte Carlo simulation for all countries that had never earned medals in the past, and the specific method is as follows. For each country i , its performance level Perf_i in each event follows a normal distribution:

$$\text{Perf}_i \sim \mathcal{N}(\mu_i, \sigma_i^2) \quad (13)$$

To account for the competition for medals, suppose there are n countries participating in event E , each with performance level Perf_i following a normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$. We perform N simulation draws. In each draw, for each country i (including smaller countries j), we sample from the normal distribution $\mathcal{N}(\mu_i, \sigma_i^2)$ to obtain the performance level Perf_i for each country.

By comparing the performance levels of all countries, we determine the ranking of each country in event E . If a smaller country j ranks in the top three, it is considered to have won a medal.

We then count the number of times country j wins a medal in N simulations, denoted as N_j , and calculate the probability that country j wins a medal in event E :

$$P_j(E) = \frac{N_j}{N} \quad (14)$$

For other possible events E_1, E_2, \dots, E_m in which country j might win a medal, we repeat the process above to calculate the medal-winning probability for country j in each event E_k . Then, the total probability of country j winning a medal across all possible events can be computed by a weighted average:

$$P_j(\text{medal}) = \sum_{k=1}^m w_k P_j(E_k) \quad (15)$$

Using the multiplication rule of probability, we calculated the probability of the number of countries that have never won a medal winning one, as shown in the **Table 3** below:

Table 3: Probability of The Number of Countries that Have Never Won A Medal Winning One

Number	0	1	2	3	4	5	6	7	8	9
P	0.06	0.18	0.25	0.22	0.14	0.08	0.03	0.01	0.01	0.00

The expected number of countries that have never won a medal to win one is 2.67. According to probability theory, we believe that two smaller countries are most likely to win medals in 2028, with a probability of 0.24.

6.3 Task 1.3

Since we use a bottom-up approach by obtaining the performance values for each country in each event every year, the ranking of each country in each event is determined by its performance value $Perf_{ij}$, where a smaller ranking indicates better performance. Let R_{ij} denote the ranking of country C_i in event E_j , which satisfies: $R_{ij} = \text{Rank}(Perf_{ij})$. $\text{Rank}(Perf_{ij})$ is the rank of $Perf_{ij}$ among all countries.

To standardize the performance values across different events, we first normalize the performance values for each event. The normalized performance value $Perf'_{ij}$ is given by:

$$Perf'_{ij} = \frac{Perf_{ij} - \min_j Perf_{ij}}{\max_j Perf_{ij} - \min_j Perf_{ij}} \quad (16)$$

This ensures that the performance values of all events are within the same range.

For each country C_i and sport $Sports_k$, which consists of multiple events, the overall performance of country C_i in sport $Sports_k$ is the average performance of all events within the sport:

$$Perf_i^{Sports_k} = \frac{1}{q} \sum_{j=1}^q Perf'_{ijk} \quad (17)$$

where q is the number of events in sport $Sports_k$.



If $Perf_i^{Sports_k}$ is larger, it indicates that country C_i has a better average ranking in that sport, implying that the sport is more important to the country.

We have created a heatmap to visually represent the process of identifying the most important sports for various countries, as shown in **Figure 12**:

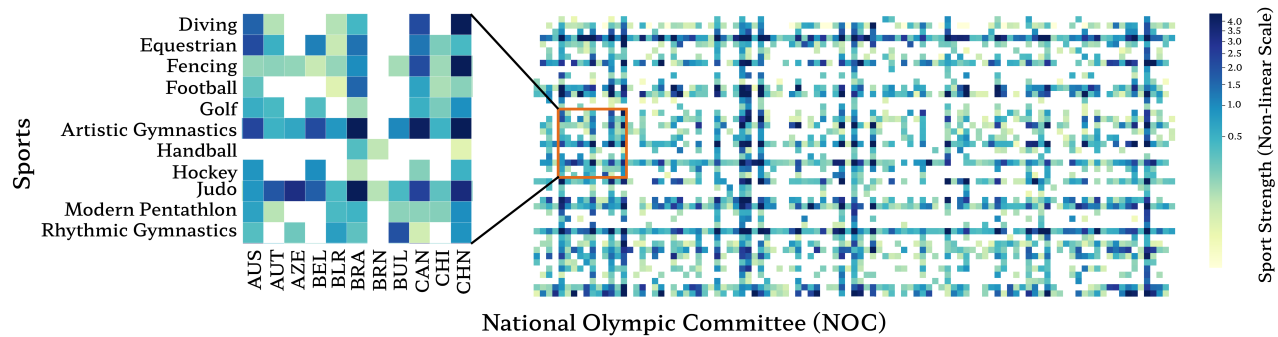


Figure 12: Identifying The Most Important Sport for Various Countries

According to the random forest model in Section 5.3, we have estimated the number of medals the host country is likely to win in newly added events. Additionally, our model found that host countries tend to perform better in these new events. The number of medals won by the host country in newly introduced events is significantly higher than that of other countries. By comparing with past data, it is also evident that the proportion of medals won by host countries in newly introduced events is much higher than that of other countries.

We visualized this phenomenon, as shown in **Figure 13** and **Figure 14** below, where red represents the proportion of total medals held by the host country, purple represents other countries, and the other colors represent the countries ranked higher on the medal table.

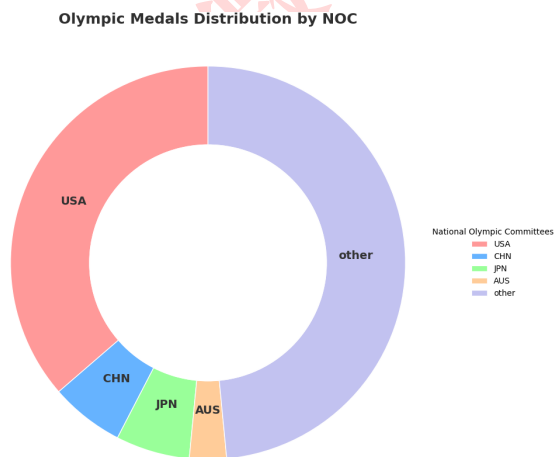


Figure 13: Proportion of Medals Won By the Host Country and Other Countries In New Events

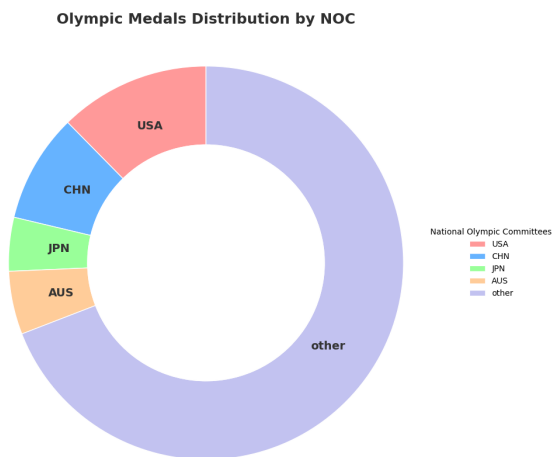


Figure 14: Proportion of Medals Won By The Host Country and Other Countries in Traditional Events

The inclusion of new sports or the re-emphasis of certain disciplines will likely boost the medal count for host nations in those typical areas.

7 Task 2: Great Coach Effect

In addition to the athletes' own performance, coaches may also influence the allocation of medals. In this task, we will focus on an unknown variable: the coach. We need to accomplish the following tasks:

- Find evidence of how coaches influence athletes' performance and medal counts.
- Analyze the impact of coach effects on medal counts.
- Conduct an analysis for three different countries to assess in which sports they should hire a great coach.

Figure 15 below shows the overall structure of Task 2:

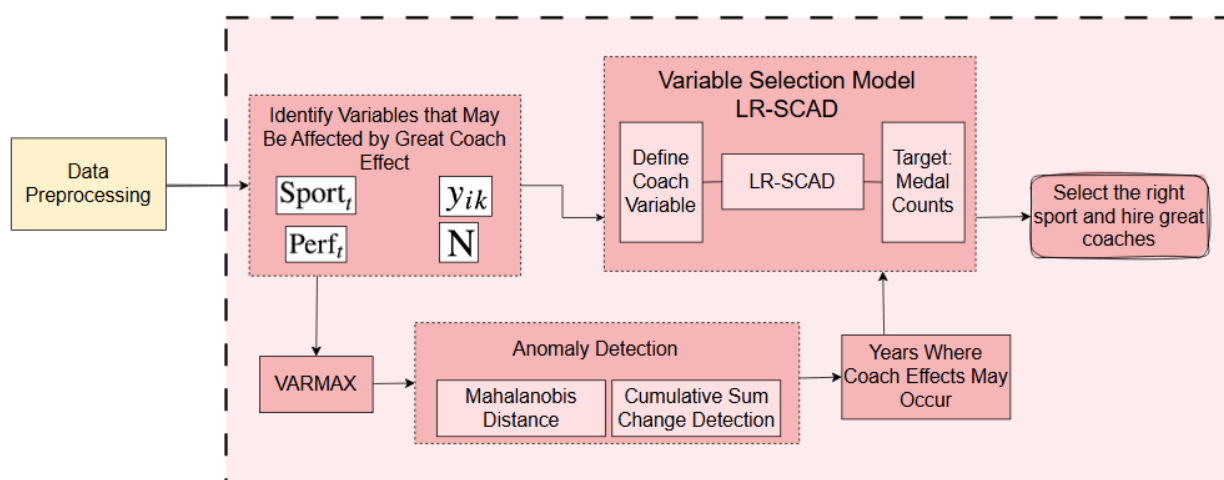


Figure 15: Overall Structure of Task 2

7.1 Residuals Test: Evidence of Changes That Might Be Due to Great Coach Effect

To identify which countries have evidence of great coaches in certain events, we implement a **residuals test**.

First, we analyze the variables most likely to be influenced by the coach effect. The following symbols all represent a specific NOC. Four variables were selected:

- Event Performance $Perf_t$: This is the yearly event performance measured in section 5.



关注数学模型
获取更多资讯

- Sport Performance $Sport_t$: The performance level in each sport, calculated in section 5.
- Number of participants N : The number of athletes participating in an event.
- Medal count y_{ik} : Medal count of country i in event k .

The core principle of anomaly year detection is to identify outliers through time series modeling combined with residual analysis.

First, the VARMAX model is introduced to perform maximum likelihood estimation to fit the time series data and solve for the model parameters[6], capturing the dynamic relationships between multiple variables.

The model is expressed as:

$$Y_t = A_1 Y_{t-1} + A_2 Y_{t-2} + \cdots + A_p Y_{t-p} + B X_t + \epsilon_t \quad (18)$$

where: Y_t is the vector of response variables, A_i is the autoregressive coefficient matrix, X_t is the exogenous variable, as is shown above, B is the coefficient matrix for the exogenous variable, ϵ_t is the white noise error term.

Next, two anomaly detection methods are introduced. The Mahalanobis distance is used to calculate the standardized residual distance, considering the covariance between variables, with the following formula:

$$D_t^2 = (\epsilon_t^\top \Sigma^{-1} \epsilon_t) \quad (19)$$

where: ϵ_t is the residual vector, Σ^{-1} is the inverse of the covariance matrix, D_t^2 is the square of the Mahalanobis distance.

Robust covariance estimation is used to reduce the impact of outliers. If the computed Mahalanobis distance exceeds the .95 percentile, it is considered an outlier.

The second method is CUSUM, which detects anomalies by calculating the cumulative sum of squared residuals:

$$S_t = \sum_{i=1}^t (\epsilon_i^2 - \mu) \quad (20)$$

where: ϵ_i^2 is the squared residual, μ is the mean of the residuals.

If the cumulative sum S_t exceeds 3 times the standard deviation, a shift point is detected, and the point is considered an anomaly.

Finally, the joint decision-making process involves taking the union of the two methods to avoid missing outliers. Anomalous points are determined by the results of both the Mahalanobis distance and CUSUM detection. The anomaly information we obtained is shown in **Figure 16** below:

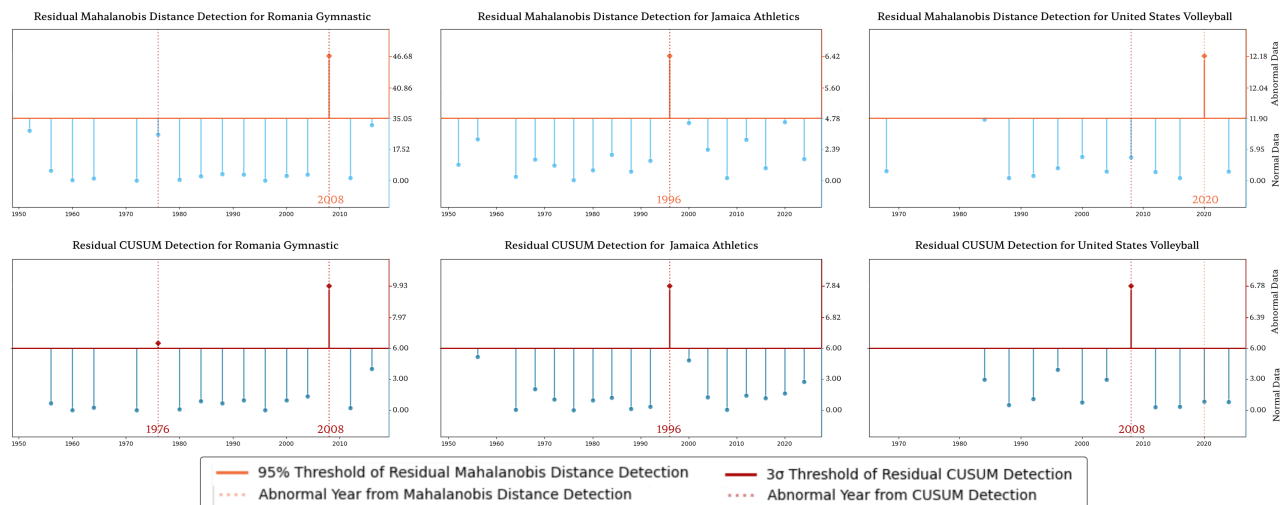


Figure 16: Anomaly Information

We found that the coaches mentioned in the problem, Lang Ping and Béla Károlyi, had data anomalies in the years they coached, confirming our hypothesis.

7.2 LR-SCAD: Estimate Great Coach Effect Contributes To Medal Counts

Next, based on the results of the residual test, we list a set of coach effect variables. Coach_{ik} is the coach effect variable for event k in country i . These variables represent the coach effect across different sports. To verify whether the coach effect significantly impacts the medal count, we include multiple potential variables affecting the medal count and the coach effect variables as independent variables in the LR-SCAD model[7].

The goal of the LR-SCAD model is to select key independent variables that affect the medal count and assess their impact. The regression equation is set as follows:

$$y_{ik} = \beta_0 + \beta_1 \cdot X_{ik1} + \beta_2 \cdot X_{ik2} + \cdots + \beta_m \cdot X_{ikm} + \beta_{\text{coach}} \cdot \text{Coach}_{ik} + \epsilon_{ik} \quad (21)$$

where y_{ik} is the medal count for country i in event k . $X_{ik1}, X_{ik2}, \dots, X_{ikm}$ are other variables that may influence the medal count. ϵ_{ik} is the error term.

After performing the SCAD model selection, if the coefficient β_{coach} is still significant (i.e., non-zero) in the regression model, it indicates that the coach effect has an independent impact on the medal count. This means that even when controlling for other variables, the coach effect still plays an irreplaceable role and cannot be explained by other variables. This result shows that the coach's role in certain sports (e.g., team sports) is crucial, and even if athletes are highly skilled, a lack of the right coach will prevent them from achieving good results.

By analyzing the SCAD parameters from the model, we can determine the specific contribution of the coach effect to the medal count for each country in each sport. We then sum these contributions to find the coach effect's contribution to the total medal count for the country:



$$\text{Contribution} = \sum_{k=1}^K \beta_{\text{coach},k} \cdot \text{Coach}_{ik} \quad (22)$$

where $\beta_{\text{coach},k}$ is the coach effect coefficient for event k .

Thus, by summing the contributions from various sports, we can determine the total contribution of the great coach effect to the country's overall medal count, as shown in **Figure 17** below:

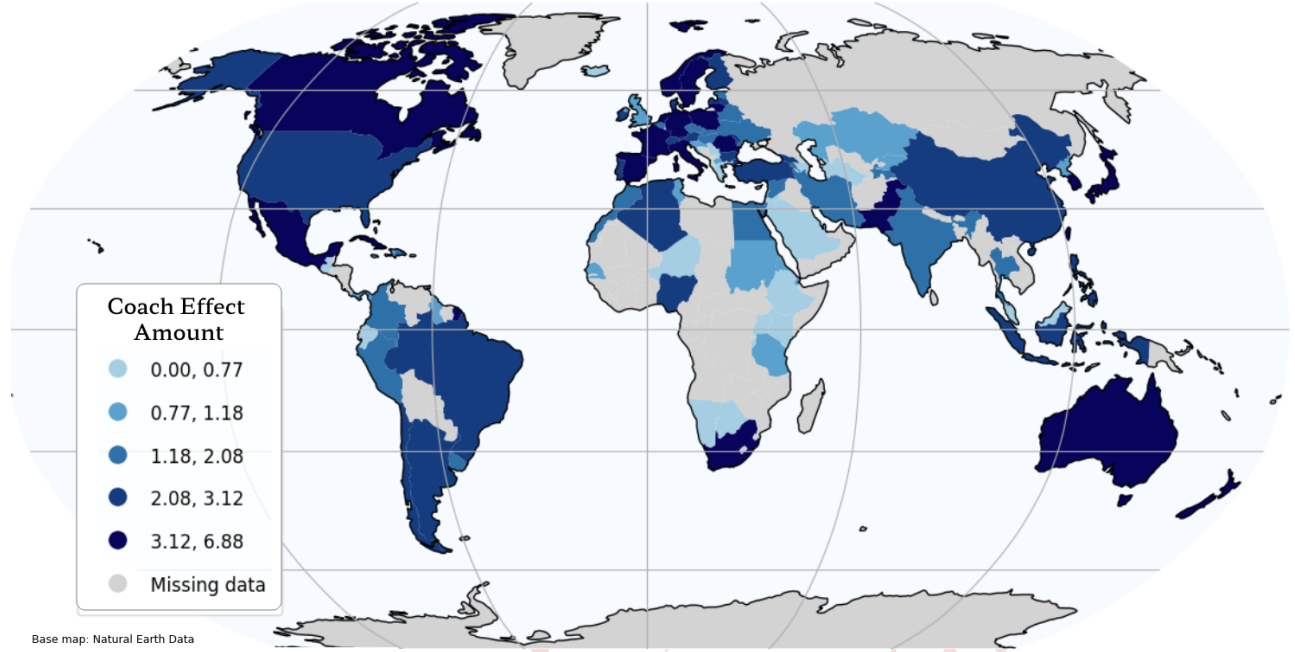


Figure 17: Great Coach Effect To The Country's Overall Medal Count[8]

7.3 Choose three countries and identify sports where they should consider investing in a “great” coach and estimate that impact

Next, we selected three NOCs, namely USA, SWE, and CHN, and calculated the great coach effect contribution for each sport. We chose the one that brings the highest benefit and calculated the resulting change in medal count, as shown in **Figure 18** below:

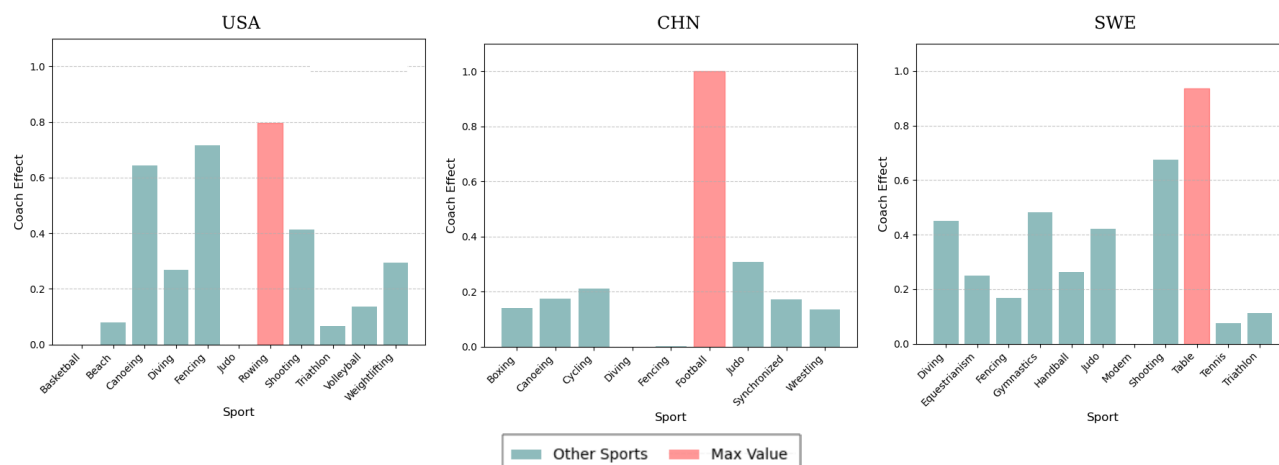


Figure 18: Great Coach Effect on Different Sports

In this regard, we made the following recommendations:

- For the USA NOC, we suggested hiring a great coach for Rowing, Fencing, and Canoeing. With the influence of a great coach, these sports are expected to add 0.8 medals to the medal table.
- For the SWE NOC, we recommended hiring a great coach for Table Tennis, expecting an increase of 0.9 medals in the medal table.
- For the CHN NOC, we advised hiring a great coach for Football, expecting an increase of 1 medal in the medal table.

In conclusion, these changes demonstrate that our model effectively captures potential scenarios involving the coach, analyzes the changes it brings to the variables, and incorporates these changes back into the model, which led to the expected impact of the coach on the medal count.

8 TASK 3: Original Insights

In this section, we will focus on original insights that were not addressed in the task but were proposed through the analysis of our model results. Based on these original insights, we will also provide effective recommendations for the Olympic Committees of the relevant countries.

8.1 Insights and Recommendations for Countries That Have Never Won Medals

Based on Section 6.2, our model identified several events in which countries that have never won medals are most likely to win. It also quantitatively shows the contribution of the coach effect to a country's medal count. This means that for countries that have never won a medal, hiring an excellent coach for the most likely events can significantly increase the probability of winning a medal.



关注数学模型
获取更多资讯

To investigate this specific impact, we applied our model to two NOCs that have never won any medals before: BEN NOC and BOL NOC. Their predicted probabilities of winning medals in the following events are illustrated in **Figure 19** and **Figure 20** below:

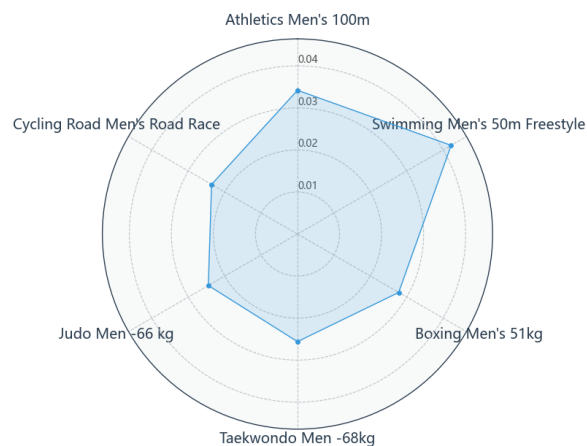


Figure 19: predicted probabilities of winning medals for BEN

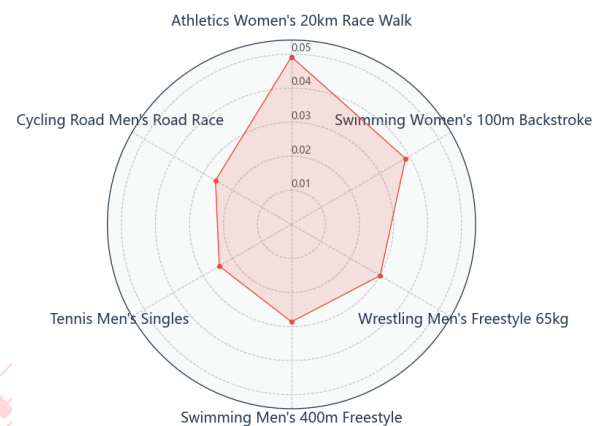


Figure 20: predicted probabilities of winning medals for BOL

Based on this, we suggest that BEN NOC and BOL NOC should adopt our model and hire Great Coaches for Swimming Men's 50m Freestyle and Athletics Women's 20km Race Walk respectively, in order to pursue their first historical medals.

Moreover, we recommend that the Olympic Committees of countries that have not yet won medals refer to our model and, taking their own strengths into account, look for excellent coaches in specific events to increase the probability of winning their first medal.

8.2 Insights and Advice for NOCs on Deploying Experienced Athletes

We also found that experienced athletes in certain events in specific countries have a significant impact on the overall level, which in turn influences the total medal count. Our model measures the country's event performance level in the sport using the Trueskill system, where the experience of athletes influences the level. Below are the overall medal statistics for USA NOC in the Fencing Men's Foil Team, the proportion of veteran athletes, and the level changes. These changes are shown in **Figure 21** below:

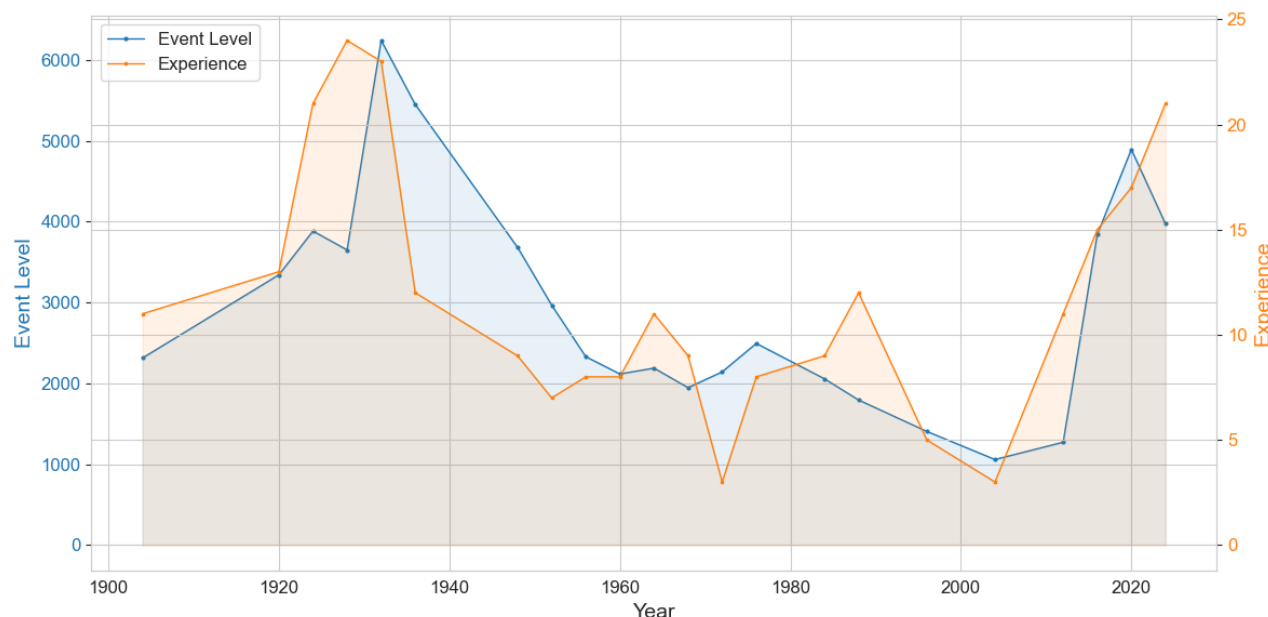


Figure 21: USA Fencing Men's Foil Team Evolution of Event Level and Athletes' Experience Over Time

This means that for USA NOC in the Fencing Men's Foil Team, having more experienced athletes participating in the event may increase the likelihood of winning a medal, compared to sending more new athletes.

Therefore, we suggest that the Olympic Committees of different countries use our model to analyze past event performance and, based on the analysis, send more experienced athletes to the competition to improve their chances of winning medals.

9 Model Evaluation

Strengths:

- We innovatively employed a variety of methods to comprehensively integrate as much useful information as possible, achieving a relatively high prediction accuracy.
- Compared to directly predicting using medal counts, the model compensate for missing cases.
- When a country doesn't participate in an event in a certain year, its level doesn't reset to zero and can still be used in subsequent predictions.
- We innovatively proposed the LR-SCAD model to accurately measure the coaching effect, which holds significant guiding value for Olympic competitions.

Weaknesses:

To ensure accuracy, a relatively large amount of computation is inevitably required. In the future, we can further optimize the model to reduce the computational cost.



关注数学模型
获取更多资讯

10 Sensitivity Analysis

In the sensitivity analysis, we selected two hyperparameters from the TrueSkill model and two from the SARIMAX model that significantly influence the results. The following **Figure 22** illustrates the variation in final medal prediction accuracy with changes in σ and β for TrueSkill, as well as p and q for SARIMAX, within a specified range.

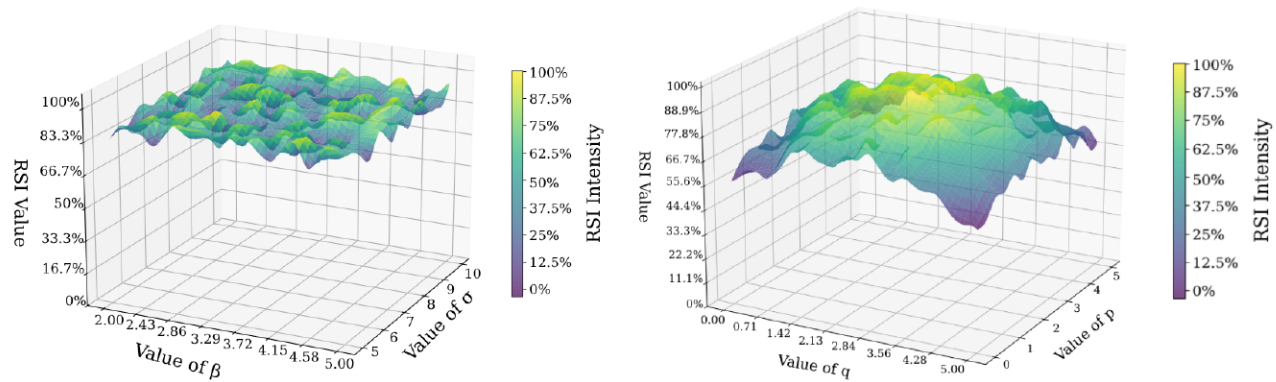


Figure 22: Variation in Final Medal Prediction Accuracy with Changes

It can be observed that when σ and β change, the prediction accuracy remains relatively stable overall, demonstrating the strong robustness of our model. This is because TrueSkill's influence on the level of each event is nearly uniform. However, prediction accuracy exhibits slightly larger fluctuations with changes in p and q . This is because these two hyperparameters directly affect the prediction outcomes.

The Following **Figure 23** shows that as the average number of samples used for time series prediction decreases, the final prediction accuracy initially tends to decline.

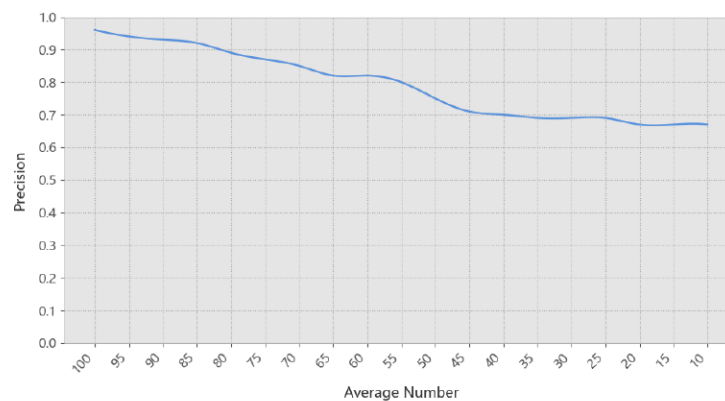


Figure 23: Correlation Between Prediction Accuracy and Sample Average Number

However, after a certain point, the downward trend slows. This is because our model automatically transitions to a simple linear regression prediction when there is less time series data, thereby avoiding larger errors introduced by time series forecasting. As a result, our model demonstrates strong adaptability across different datasets.

11 Conclusion

For **Task 1**, we built a bottom-up medal analysis and prediction model using the TrueSkill-SARIMAX-RF method. The model first categorizes all events into traditional events and new events. For traditional events, the TrueSkill model is used to evaluate a country's performance in each event. Then, the SARIMAX model is applied to predict the country's performance level for 2028. To account for potential errors, a Bayesian prediction method is employed. Finally, the predicted performance for 2028 determines the medal allocation. For new events, a Random Forest model is used to predict medal allocation. The results from these two parts of events are aggregated to form the final medal count model for each country.

For **Task 2**, we performed residual analysis to find evidence of how coaches influence athletes' performance and medal counts. By analyzing different variables, we identified anomaly information in the data, providing evidence of how the "great coach" effect alters the data. Subsequently, we studied how the "great coach" effect contributed to changes in the total medal count for different countries and created visualizations for this. Finally, we selected USA NOC, SWE NOC, and CHN NOC and identified the sports in which these countries should invest in a great coach. We then evaluated the impact of the great coach on their medal counts.

For **Task 3**, based on the model, we analyzed and derived 2 factors that affect a country's medal count, which were not covered in Tasks 1 and 2. These factors were then integrated into insights and explained how these insights could provide valuable reference for national Olympic committees.



References

- [1] Wikipedia contributors. Wikipedia. <https://en.wikipedia.org/>. Retrieved January 28, 2025.
- [2] T. Minka, R. Cleven, and Y. Zaykov. Trueskill 2: An improved bayesian skill rating system. Technical report, Microsoft Research, n.d.
- [3] W. Skaf, A. Tosayeva, and D. T. Várkonyi. Towards automatic forecasting: Evaluation of time-series forecasting models for chickenpox cases estimation in hungary. *arXiv preprint arXiv:2209.14129*, 2022.
- [4] P. Berti, E. Dreassi, F. Leisen, P. Rigo, and L. Pratelli. A probabilistic view on predictive constructions for bayesian learning. *arXiv preprint arXiv:2208.06785*, 2022.
- [5] A. Sinha and S. Bhatia. Deep regression forests. In *Proceedings of the IEEE International Conference on Computer Vision (ICCV)*, pages 3653–3662, 2019.
- [6] Deepvarma: A hybrid deep learning and varma model for chemical industry index forecasting. No additional information provided on authorship or publication year.
- [7] J. Ma, L. Wang, and T. Xu. Approximate message passing for nonconvex sparse regularization with stability and asymptotic analysis. *arXiv preprint arXiv:1711.02795*, 2017.
- [8] Natural Earth. 1:110m cultural vectors. <https://www.naturalearthdata.com/downloads/110m-cultural-vectors/>. Retrieved January 28, 2025.