Problem Chosen	2025	Team Control Number
C	MCM/ICM Summary Sheet	25*****

Here Be Prophets: Unraveling the 2028 LA Olympic Games

Summary

The quote by Tagore, "Life is in sports," highlights the importance of sports. As an international sporting event, the Olympic Games have attracted much attention. In this paper, we focused on the number of medals won by each country in the Olympics and developed a systematic analysis to provide insights for the decision-making of National Olympic Committees.

Before constructing the model, we assessed the data quality, integrated and cleaned the data. Considering the condition "use only the provided data sets," we applied **Wilcoxon rank sum test** to prove the significant impact of gender distribution on medal outcomes, supporting the subsequent modeling process.

First, we constructed features such as the number of participants in specific events CS_pants. Through **Levene test**, we found that the features did not possess homogeneity of variance. So we chose **Multi-Layer Perceptron (MLP)**, which does not have distribution requirements. Experiments showed that on the test set, the MLP achieved an R^2 of **0.9425** for total medals and **0.7654** for gold medals, outperforming other algorithms (see Figure 6). We further applied MLP to predict the 2028 USA Summer Olympics, with visualized diagrams, where we, by reasoning, concluded that Samoa ,etc. may win their first medal.

Secondly, now that our data are sequences and did not possess homogeneity of variance, we used **Gray Relational Analysis (GRA)** to validate the significant correlation between events and medals won by the country. And to identify the most important sports for each country while combining objectivity and subjective preferences of each country, we innovatively minimized the risks of mis-classification quantified by KL divergence, thereby deriving composite weight values. Then we developed the **KL-CEAHP-VIKOR** evaluation model with the composite weight values. This model identified Athletics as the most important sport for the USA, in line with the reality. Additionally, the results of the previous prediction model and evaluation model indicated that the selection of events will have a significant influence on the distribution of medals.

Furthermore, to investigate the "great coach" effect, we first identified the USA, China, and the Netherlands as research subjects. The **Moving T Test (MTT)** confirmed the presence of the "great coach" effect, and the **Granger causality test** showed that the effect does not influence other events led by non-foreign coaches. Finally, we quantified the contribution of the "great coach" effect with **Interrupted Time Series Analysis (ITSA)**, and illustrated its gain and investment potentials.

Finally, based on the model results, we derived several insights and provide four recommendations for National Olympic Committees (NOCs). We also performed an evaluation of merits and demerits, and **sensitivity analysis** on our models, demonstrating their robustness.

Keywords: Wilcoxon rank-sum test, MLP, KL-CEAHP-VIKOR, ITSA

Contents

1 Introduction

1.1 Problem Background

The Olympic Games are a global event held every four years, uniting athletes and promoting excellence and camaraderie. Beyond watching athletes' performance, the audience also keeps up with the overall "medal table". Front-runners in the table naturally gain much attention, but other countries are worth seeing as well, such as those winning a medal for the first time. From this, predictions of the winner of each medal and the final medal counts are generally made by both sports fans and professional institutes.

1.2 Restatement of the Problem

Conventionally, predictions of medal counts are made based on knowledge of the athletes scheduled to compete, rather than on historical medal counts.

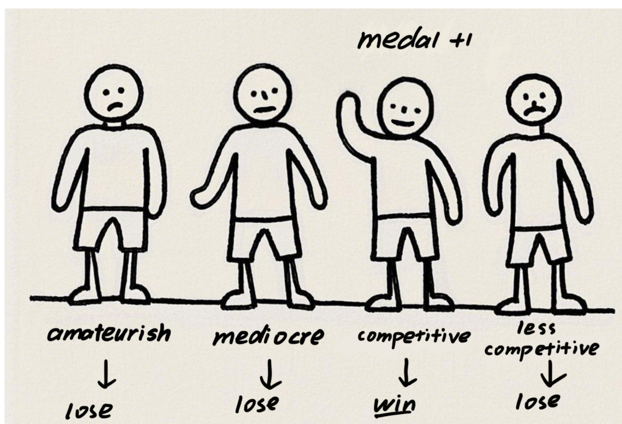


Figure 1: Athlete-Based Prediction

Accordingly, we are required to employ the given historical data sets to build mathematical models and solve the following problems:

- Forecast minimum medal counts for each country (gold and total medals), assess model uncertainty and precision, and define performance metrics.
  - Project the Olympic medal table for Los Angeles 2028, include prediction intervals and identify countries most likely to improve or worsen.
  - Estimate the number of countries yet to win medals that will earn their first in the next Olympics, and provide the odds.
  - Analyze how Olympic events affect medal counts, determine key sports for each country, and explore how host country sports impact medal tallies.
- Investigate the impact of the **"great coach" effect** on medal counts; select three countries to identify sports where this effect should be applied and estimate its impact.
- Extract insights into medal counts from our model and provide recommendations to Olympic committees.

1.3 Our work

The flow chart of our work is shown in Figure 2 below:

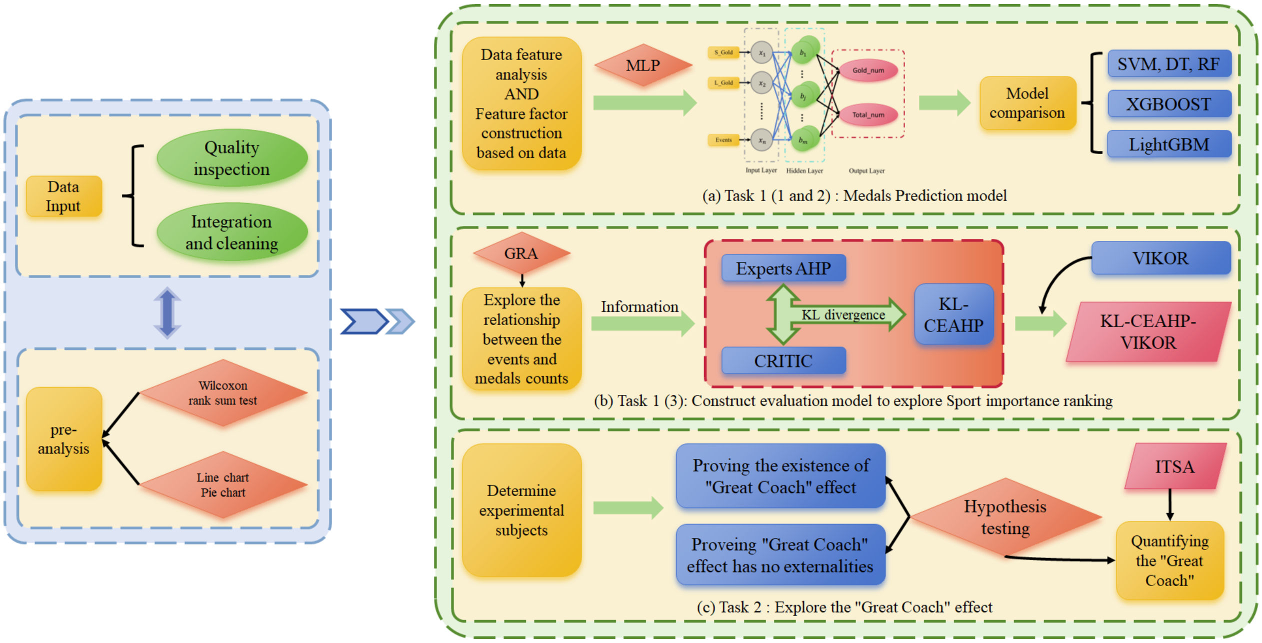


Figure 2: Our Work

2 Assumptions and Justifications

We made several general assumptions to simplify the model. These assumptions and the corresponding justifications are as follows:

1. **The athlete data includes all participants in the current Olympic Games.**
   To predict a country's medal count, it is essential to analyze the athletes representing that country, including their event participation and the total number of athletes involved. Thus, the completeness and validity of relevant data should be guaranteed.

2. **Other events, excluding major political or financial crises, have negligible impact on the Olympics.**
   Minor events do not disrupt the Olympics. For model simplicity, such factors are not considered.

3. **The host country has subjective preferences for certain sports.**
   Hosts often promote sports that reflect their unique culture, influencing the selection of events.

4. **National Olympic Committees (NOCs) make rational decisions.**
   NOCs allocate resources based on their country's conditions and local preferences.

3 Notations

The primary notations used in this paper are listed in Table 1.

Table 1: Notations

Symbol	Description
U	Statistical Test Statistic
t	Normalized Time Stamp
$y(t), y_t$	Experimental Goal Sequence
$x_i(t), x_t$	Explanatory Variable Sequence
W	Mixed Subjective and Objective Weighting
S, R, Q	Group Utility Value, Individual Regret Value, and Decision Indicators in the VIKOR Method

4 Data Pre-processing

4.1 Data Cleaning

Due to the limitation of using only the provided dataset, it is essential to investigate the data quality and perform data cleaning. The results of this process are outlined below:

- No missing values in the key data (athlete data).
- No duplicate entries in the key data (athlete data).
- No significant outliers in the key data (athlete data).

However, during the analysis process, we identified some significant issues related to information fragmentation:

- In the "summerOly_medal_counts.csv" dataset, there are instances of data entries such as "<0xa0>" For example, the country "United States" appears in two forms: "United States" and "United States<0xa0>." These entries need to be merged.
- In the "summerOly_athletes.csv" dataset, using Volleyball as an example, the "Event" column contains values such as "Volleyball Women's Volleyball", "Women Team", and "Women", all of which refer to women's volleyball, but are fragmented into three different expressions. Similarly, the "Team" column also contains multiple representations of the same country, which necessitates the mergence of relevant data.

4.2 Analysis of Influential Factors

Previous medal count predictions were typically based on athlete data, and we have similar information in "summerOly_athletes.csv". Therefore, it's essential to test whether and how the other variables in "summerOly_athletes.csv" affect the "Medal" variable. This variable denotes the type of medal(gold, silver, bronze or no medal).

First, for elite athletes, the prime competitive age range mainly falls between 22 and 30 years old, while a minority of athletes may have a slightly longer competitive age range[1]. Thus, the 2024, 2020 and 2016 Olympic Games are contributive to the forecast of the 2028 Olympic Games. In the following analysis, unless otherwise specified, a time span of these three years(2024, 2020 and 2016) will be used for research, to weaken the horizontal influence of various time spans.

4.2.1 Impacts of Gender

Gender, culture, and athlete build all influence competitive sports outcomes. We first explore gender's impact on the "Medal" variable.

To isolate other factors like culture, we document the gold, silver, bronze medals, and non-medalists (assigned 0 if none) separately for Male and Female athletes in each sport. These counts are then normalized into percentages to eliminate scale disparities. The data for Male and Female athletes are sorted by sport, yielding two comparable datasets.

Since the data do not follow a specific distribution, traditional methods including ANOVA(Analysis of Variance), which rely on data distribution and variance homogeneity, are not applicable. Therefore, we use the non-parametric Wilcoxon rank sum test, which does not assume any specific distribution, to determine if gender significantly impacts the "Medal" variable.

For the USA, with a significance level $\alpha = 0.1$, the Wilcoxon rank sum test statistic is constructed as follows:

1. Let the distribution of the "Medal" variable among male athletes in the USA be Ω_a, and its median be m_a. For female athletes in the USA, let the corresponding distribution and median be Ω_b and m_b respectively. We propose the following test hypotheses:
$$\begin{cases} H_0 : \Omega_a = \Omega_b, m_a = m_b \\ H_1 : \Omega_a \neq \Omega_b, m_a \neq m_b \end{cases}$$

2. Combine all the elements from the two samples and sort them in ascending order, where n_a and n_b denote the sample sizes of the two samples respectively. Let $x_{c(i)}$ be the sequence derived from sorting the combined $x_{a(i)}$ and $x_{b(i)}$
$$x_{c(1)} \leq x_{c(2)} \leq \cdots \leq x_{c(n_a+n_b)},$$
$$x_{c(i)} \in \mathrm{sort}(\{x_{a(1)}, x_{a(2)}, \ldots, x_{a(n_a)}, x_{b(1)}, x_{b(2)}, \ldots, x_{b(n_b)}\})$$

3. Let R_a be the rank sum of all $x_{a(i)}$(i.e., the sum of the ranks of elements in $x_{c(i)}$ that are derived from $x_{a(i)}$). Similarly, we have R_b. Moreover, we introduce complementary

rank sum R_a':

$$R_a' = na \times (na + nb + 1) - R_a$$

Then, we have Wilcoxon rank sum test statistic $U = \min(R_a, R_a')$.

In terms of the United States' gold medal count, the Wilcoxon rank sum test statistic U is -2.7948, the corresponding P value is 0.0052, which is far less than 0.1. As for silver medal count, bronze medal count and non-medalists, the P values are respectively 0.3152, 0.9404 and 0.9072. This indicates that for the United States, gender has a significant impact on gold medal tally with regard to competitive sports, but has no significant impact on other medal tally.

For generality, we conducted the same procedure for USA(United States of America), CHN(China), FRA(France), NED(Netherlands), GBR(the Great Britain) and RSA(the Republic of South Africa). The test statistics are presented in Table 2:

Table 2: Wilcoxon Test Statistics for Various Countries

	USA	CHN	JPN	AUS	FRA	NED	GBR
Gold	0.0052***	0.0383**	0.9949	0.7085	0.2097	0.882	0.573
Silver	0.3152	0.244	0.82	0.4765	0.3289	0.5791	0.352
Bronze	0.9404	0.6994	0.9295	0.8352	0.0795*	0.2043	0.3339
No medal	0.9072	0.3984	0.6446	0.2458	0.0009***	0.1973	0.5669

Note: ***, ** and * represent significance levels of 1%, 5% and 10% respectively.

We can clearly see from the table whether Gender has significant influences under diverse conditions.

4.2.2 Impacts of Being the Host or Not

Admittedly, the Olympics advocate for mutual understanding, friendship, unity and fair competition. Nevertheless, being the host country does facilitate its athletes. Therefore, we will explore the potential influence of being the host or not on medal counts.

Similar to 4.2.1, we will control the influence of other factors. Additionally, in the 1980s, 1990s and earlier, global economic and political upheavals directly or indirectly affect Olympic Games. Hence, we take the Olympic Games held in 2000 and later as our research objects.

Through data screening, we identified the hosts of the six Summer Olympic Games from 2000 to 2024, and recorded gold, silver, bronze and total medal counts in the year of hosting, one year before, two years before and three years before. The visualized results are shown in Figure 3.

We can conclude from the figure above that being the host or not does have positive impact on medal tally, since the medal counts(gold, silver, bronze or total) of these six countries witnessed upward trends in the year of hosting. We will prove significance and quantify the contribution to medal counts in the next section.

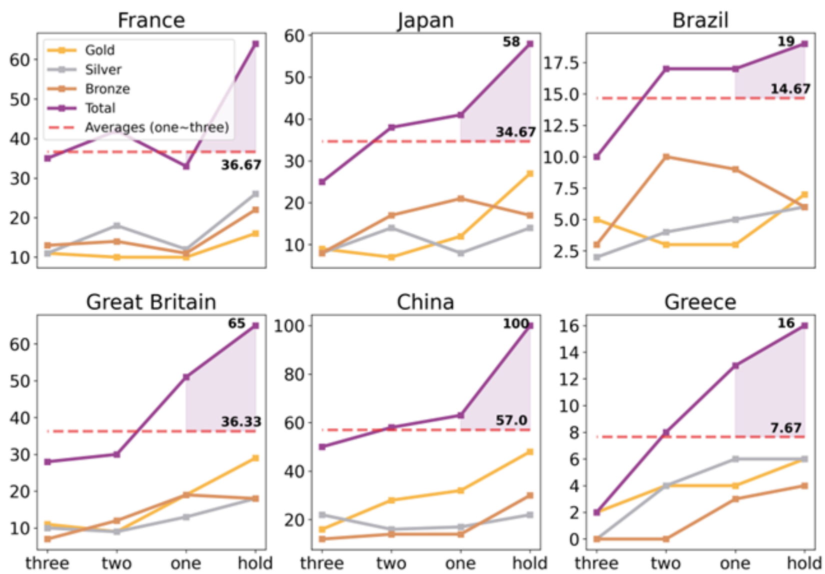


Figure 3: Medal Counts of Six Host Countries from 2000 to 2024

4.2.3 Impacts of A Country's Strengths

Apparently, athletes from different countries and regions exhibit distinct physical characteristics. And they excel in different sports disciplines. For almost any country, the number of gold, silver, and bronze medals, as well as the total medal count, is largely determined by the sports in which they have particular expertise. Therefore, accurately predicting the medal tally for each country in the next Olympic Games requires precise identification of the sports in which the country excels.

In our study, we focus on two popular sports, Athletics and Shooting, out of a total of 71 sports. We record the medal statistics (gold, silver, and bronze) for each country during the specified time period(2024, 2020 and 2016).To facilitate better visualization, we present the data as 3D pie charts below (with Athletics on top and Shooting at the bottom):

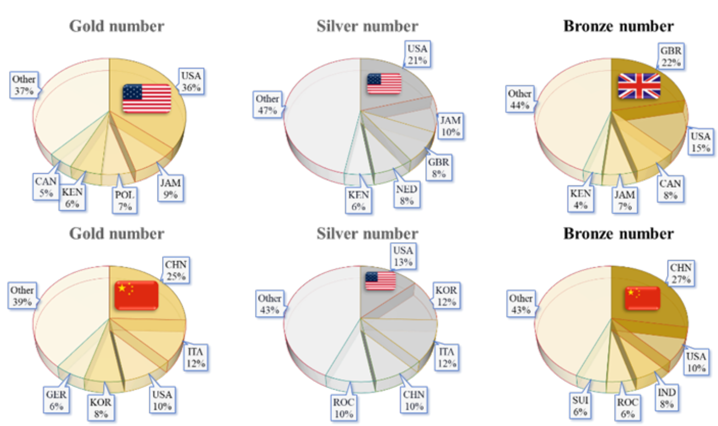


Figure 4: Medal Distribution for Athletics and Shooting by Percentage

In Athletics, the USA holds a dominant position, with an absolute advantage in both gold and silver medals, accounting for more than one-third of the gold medals. Even in the bronze medal category, the USA ranks second with a 15% share. Based on this, we have reason to consider Athletics as one of the USA's "specialized events." In the subsequent process of building prediction models or in the analysis of important competitions for the USA, Athletics should be prioritized as one of these "specialized events."

In Shooting, China demonstrates a strong performance with a leading position in both gold and bronze medals. China ranks first in gold medals, accounting for 25% of the total, and first in bronze medals with 27%. Although China ranks fifth in silver medals, with a share of 10%, its overall performance in Shooting remains dominant. Given this, we can conclude that Shooting is one of China's "specialized events." And Shooting should also be prioritized in the subsequent analysis.

5 Task 1: Predictions of the Medal Tally

5.1 Problem 1: Projections for the 2028 Olympics Medal Table

5.1.1 Feature Engineering

To predict the medal count for each country, we will consider not only historical medal data but also athlete-related information to account for the specific conditions of the current Olympic Games while excluding external factors. Therefore, we will conduct feature engineering at the event level for each country. New features are constructed as follows:

- **Short-term Gold Medal Average (Short_Gold) and Short-term Total Medal Average (Short_Total):** As indicated in 4, the optimal athletic career span is approximately ten years, namely there will be overlap in athlete participation across three consecutive Olympic Games. This provides relevant information for prediction. Thus, the average medal count of the past three Olympic Games is considered as a feature factor. For example, the formula for Short_Gold is:

$$Short\_Gold_t = \frac{1}{3}\sum_{i=t-3}^{t-1} \mathrm{Gold}_t$$

- **Long-term Gold Medal Average (Long_Gold) and Long-term Total Medal Average (Long_Total):** This feature reflects the traditional powerhouse countries in specific events, such as the USA in Athletics. The average medal count of the last ten Olympic Games is considered as a feature. For example, the formula for Long_Gold is:

$$Long\_Gold_t = \frac{1}{10}\sum_{i=t-10}^{t-1} \mathrm{Gold}_t$$

- **Country-specific Short-term Gold Medal Average (Short_CGold), Short-term Total Medal Average (Short_CTotal), Long-term Gold Average (Long_CGold) and**

Long-term Total Medal Average (Long_CTotal): The two features above reflect information for a specific event to assist in local adjustments. However, on nation-scale predictions, corresponding metrics are needed. The formulas of the metrics are the same as those above.

- **Host Country Indicator (Is_host):** As observed in Figure 3, there is a significant increase in medal counts when a country hosts the Olympic Games. Thus, a 0-1 intervention variable Is_host is introduced, where 1 represents being the host country and 0 otherwise, to account for the host effect.

- **Total Participants for Each Country (C_pants), Event-specific Participants for Each Country (CS_pants), Total Participants in an Event under a Sports (S_pants), and Number of Events in that Sport (E_num):** Clearly, a greater value of CS_pants for a given country indicates greater emphasis on that event, thus higher probability of winning medals. The other features have similar implications. Additionally, we introduce the last feature *Ratio*, which represents the proportion of a country's participants in a specific event relative to the total participants in that event at the Olympic Games:

$$Ratio = \frac{CS\_pants}{S\_pants}$$

With these features, all the implicit information in the dataset has been revealed, assisting in the construction of the subsequent predictive model.

5.1.2 Model Establishment

First, visualize the data distribution in Figure 5:

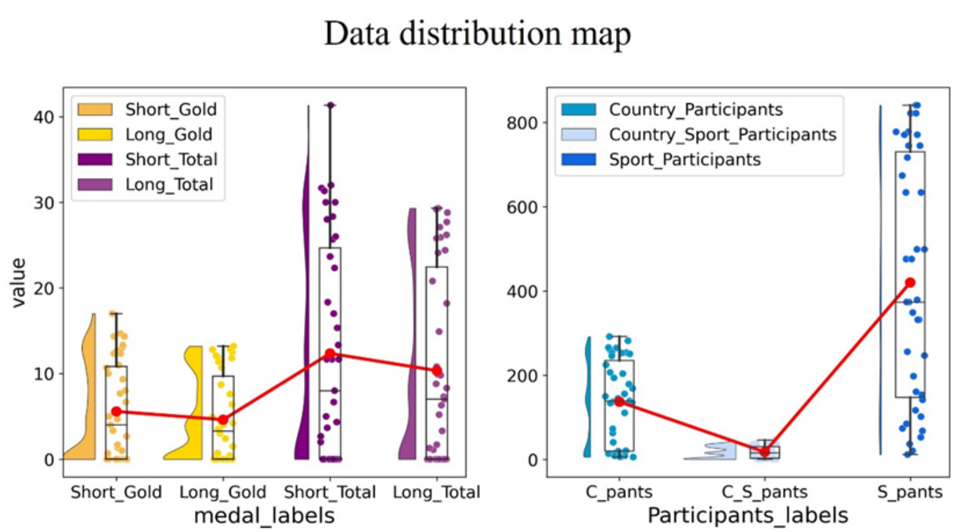


Figure 5: Data Distribution Map

We noticed that the features do not exhibit a specific distribution and do not possess homoscedasticity. Traditional statistical prediction models, such as linear regression, have strict requirements for data distribution and are not suitable for these features. Additionally, there may be implicit relationships between the features. Therefore, we consider building a multilayer perceptron (MLP) model, which does not impose specific distributional assumptions and can accommodate relatively complex relationships. Its structure is shown in the left figure below:

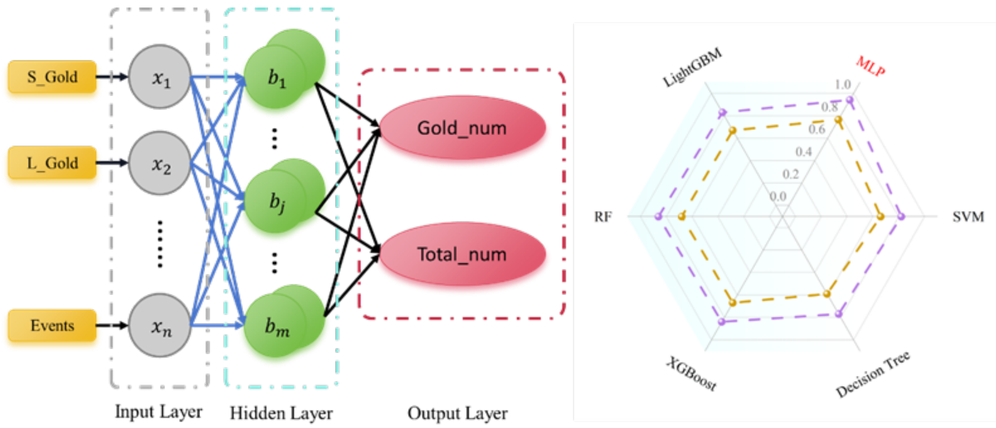


Figure 6: The Structure of MLP and Model Performance Comparison

The features and class labels are input into the MLP for training. Subsequently, we tested the model on the 2024 test set, and found that the goodness of fit R^2 for the total medal count reaches 0.9425, while the R^2 for the gold medal count is 0.7654. Additionally, it is observed in the right figure above that the goodness of fit measure R^2 of the MLP model outperforms well-known machine learning algorithms, such as XGBoost, and is better suited for this prediction task comparing to other algorithms.

5.1.3 Prediction Results

The MLP model was trained using data prior to 2024, with 2024 as the validation set. The prediction accuracy for 2024 is as follows:

- Gold Medals: RMSE = 2.400, $R^2 = 0.765$
- Total Medals: RMSE = 3.410, $R^2 = 0.942$

Considering the limited information available, an R^2 above 0.75 indicates a good capture of the relevant patterns. The model produced the projection of the 2028 Olympics medal table, which is shown below:

NOC	Country	Gold	Total	NOC	Country	Gold	Total
United States		60	101	Great Britain		17	40
China		44	71	Japan		12	35
Germany		21	41	Italy		12	28
Australia		17	54	Republic of Korea		11	18
France		17	46	Canada		6	22

Figure 7: Projection of Medal Table for 2028 Los Angeles Olympics

We also calculated 95% prediction intervals, partly shown in Table 3:

Table 3: Prediction Intervals

	NOC	Gold	G_low	G_upp	Total	T_low	T_upp
1	USA	60	55	64	101	94	107
2	CHN	44	39	48	71	64	77
3	GER	21	16	25	41	34	47
4	AUS	17	12	21	54	47	60
5	FRA	17	12	21	46	39	52

To provide a more comprehensive analysis of the results, we extracted the five countries with the greatest changes in medal counts, as is shown below in Table 4:

Table 4: Medal Count Changes

	Gold	Total		Gold	Total
USA	60(+20)	101(-25)	NED	5(-10)	21(-13)
GER	21(+9)	41(+8)	JPN	12(-8)	35(-10)
CHN	44(+4)	71(-20)	NZL	4(-6)	8(-12)
GBR	17(+3)	40(-25)	UZB	3(-5)	9(-4)
BRA	5(+2)	17(-3)	BUL	0(-3)	9(+2)

Note: The left side represents increases and the right side represents decreases.

To analyze the increase in medal counts, we use the United States as an example. Considering that the medal table ranking is based on the gold medal count, its the importance is greater than the total medal count. It can be observed that the increase in the gold medal count for the United States has resulted in a significant lead, which can be inferred from the projection for the USA being the host of the 2028 Olympics, validating the feasibility of using being host or not as a feature.

As for the decline, using Japan as an example, in the 2020 Tokyo Olympics, Japan, as the host, possibly benefited from the host effect. Therefore, the gold medal performance in the Tokyo Olympics could be considered an "outlier" and might not fully represent the long-term trend. This implies the decline in Japan's medal count for subsequent Olympics, which is consistent with the forecasted results.

5.2 Problem 2: Projections for First-Time Medalists in the 2028 Olympics

Based on previous analysis in 5.1, in the prediction of medal counts for the 2028 Summer Olympics, our model forecasts that **Samoa** is likely to win their first medals. Through data analysis, we observed that Samoa had 13 participants in the 2024 Olympics, an increase of 2 compared to 2020, which is a notable change.

Additionally, by analyzing the medal counts for the 2024 Olympics, the model successfully predicted that the EOR (formerly known as the Refugee Olympic Team) would win medals. Besides, this team had 40 participants, far exceeding other smaller countries that won medals for the first time in the same year (e.g., Saint Lucia with 5 participants). Therefore, the number of participants is one of the key factors contributing to the success of the prediction model, also confirming the likelihood of countries like Samoa winning medals for the first time.

Finally, we quantified the odds of the prediction results by defining the odds as follows, where *true_num* is the number of countries in the test set that won medals for the first time, and *pred_num* is the predicted number of such countries:

$$\text{odds} = \frac{\frac{\text{pred\_num}}{\text{true\_num}}}{1 - \frac{\text{pred\_num}}{\text{true\_num}}}$$

Substituting the data with values, the predicted odds were be 20%. Given that the medal counts are concentrated among sporting powerhouses, this is a relatively high odds (the higher the odds, the better), indicating that our model is capable of reflecting the medal outcomes of smaller sports nations.

5.3 Problem 3: Impact of Events on Medal Counts and Country Performance

1. **Your model should also consider the events (number and types) at a given Olympics.**

Our model accounts for the number and types of events at the Olympic Games, incorporating their impact on medal counts for each country.

2. **Explore the relationship between the events and how many medals countries earn.**

Herein, we aim to investigate the relationship between Olympic events and a nation's medal counts, focusing on its performance reflected in both gold medal count and total medal tally. We propose utilizing event-related features, gold medal count and total medal count, to quantify their associations through an appropriate correlation measure.

From the dataset, for each country, three event-related features are extracted:

- Number of participants per event : P_{num}
- Total number of events participated: E_{num}
- Participation ratio: Ratio $= \frac{P_{\text{num}}^{(\text{country})}}{P_{\text{num}}^{(\text{global})}}$

As demonstrated in prior analysis:

- Variables P_{num}, E_{num}, and *Ratio* exhibit no specific statistical distribution, rendering Pearson correlation unsuitable.
- Longitudinal data across multiple Olympics are required to capture temporal patterns.

- Spearman correlation is inappropriate due to its rank-ordering mechanism that erases sequential structure.

We adopt Grey Relation Analysis for sequential data association, implemented through the following steps:

$$\xi_i(k) = \frac{\underset{i}{min}\ \underset{k}{min}|y(k) - x_i(k)| + \rho \underset{i}{max}\ \underset{k}{max}|y(k) - x_i(k)|}{|y(k) - x_i(k)| + \rho \underset{i}{max}\ \underset{k}{max}|y(k) - x_i k|}$$

where:

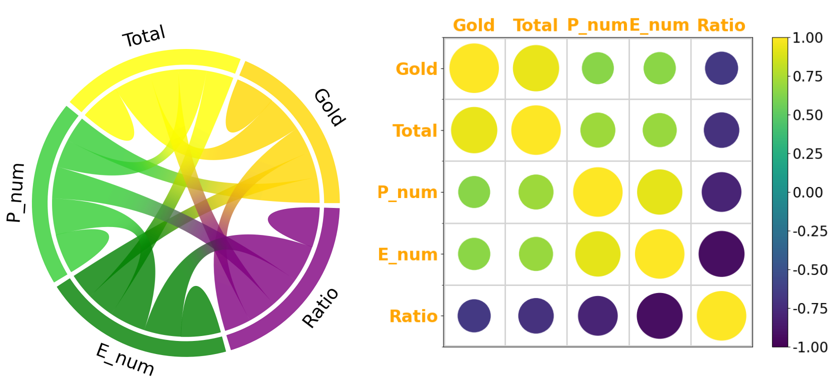
- $\xi_i(k)$: the correlation coefficient of x_i with respect to $y(k)$ at position k.
- $|y(k) - x_i(k)|$: the absolute difference between y and x_i at position k.
- $\underset{i}{min}\ \underset{k}{min}|y(k) - x_i(k)|$: the minimum absolute value of the second-order difference between sequence y and sequence x_i at position k.
- $\underset{i}{max}\ \underset{k}{max}|y(k) - x_i(k)|$: the maximum absolute value of the second-order difference between sequence y and sequence x_i at position k.
- ρ: the gray relational coefficient, ranging from 0 to 1, typically set to 0.5.

By substituting the correlation coefficients of each variable into the following equation, the correlation degree r_i between $x_i(k)$ and $y(k)$ can be computed:

$$r_i = \frac{1}{n}\sum_{k=i}^{n} \xi_i(k)$$

By comparing the geometric shape similarity between a reference data sequence and other data sequences, we determine the degree of correlation between data them.

Using the USA as an example, a gray relational analysis is conducted on its gold medal count and total medal count in relation to the historical data of P_{num}, E_{num}, and *Ratio*. The visualized results are displayed in Figure 8:



(a) Chord Diagram (b) Heat Map

Figure 8: Correlation Analysis

The left panel presents a chord diagram that visually represents inter-feature correlations through regional connections and bandwidth variations of linking arcs. The right panel displays a heat map that more explicitly quantifies both statistical significance and magnitude of these correlations.

Analytically, a nation's level of participation — both the number of participants per event and the total number of events within the same category of sports — reflects the breadth and depth of its involvement. These participation metrics serve as a proxy for national prioritization of specific sports. Quantitative accumulation leading to qualitative transformation, such comprehensive participation statistically enhances medal acquisition probabilities. This aligns with the observed strong positive correlations between gold/total medal counts and these participation metrics, as evidenced by the USA dataset.

The negative correlation between *Ratio* and gold/total medal counts warrants particular attention. This phenomenon could be attributed to differential growth rates - where global participation expansion in certain events has outpaced corresponding US engagement growth. However, newly participating nations' athletes generally demonstrate less competitiveness compared with established sporting powers, thereby not constituting substantial competitive pressure.

The same procedure conducted on other countries produced similar results. This consistent observation strongly supports the associations between event participation levels and national medal achievements.

3. **What sports are most important for various countries?**

As demonstrated in 4.2.3, different countries exhibit variations in sports specialization. As illustrated in Figure 4, the USA dominates in Athletics, ranking first in both gold and silver medals, and second in bronze medals. Clearly, Athletics can be considered one of the USA's key sports. In contrast, China does not rank among the top in any medal category, suggesting that Athletics is probably not a key sport for China. However, in Shooting, while the USA remains strong, China leads in both gold and bronze medals. Therefore, although Shooting can still be considered one of the USA's important sports, China holds a significant advantage and can also be regarded as having Shooting as one of its key sports.

Naturally, subjective observations lack generalizability. Hence, we consider using the relevant features we constructed to build a quantification model, providing a more objective approach to identifying the key sports for each country.

In this section, we consider that the number of medals won in a sport and athlete deployment intensity may reflect the country's emphasis on that sport. Therefore, we constructed six features indexed by sports: "Short_Term_Item_Gold", "Long_Term_Item_Gold", "Short_Term_Item_Total", "Long_Term_Item_Total", "Ratio" and "Events" as evaluation factors.

First, we estimate the weights of these six evaluation factors in the model. For objectivity, we apply the CRITIC(Criteria Importance Through Inter-criteria Correlation) method here. This method determines the objective weight of each evaluation factor based on the factor's contrast intensity and the conflicts between factors. It incorporates the variability of the factors while accounting for their correlations, making it more objective and reliable than the entropy weighting method or standard deviation weighting method.

Next, given that subjective experience is also an important factor in weighting for evaluating the importance of a sport, we also incorporate the Experts AHP (EAHP) model, which more accurately reflects a country's real-world preferences.

Furthermore, we discovered that hybrid weighting approaches, such as AHP-entropy, AHP-OWA, and others, are prevalent in evaluation model frameworks. These methods typically directly apply weights to construct new weights. However, this approach does not consider the coherence between original weights nor the information envelope between new and original weights. Therefore, this hybrid weighting method is not suitable.

Consequently, we turn to the Kullback-Leibler divergence (KL divergence), which measures the difference in distributions, and adopt its framework to construct envelope discriminant information between new weights and original weights. The discriminant information as the objective, we minimize the information difference between the new weights and original weights, employing the method of Lagrange multiplier to derive the composite weight values:

The CRITIC method yields the weight vector $\mathbf{W}^A = (w_1^A, w_2^A, \ldots, w_n^A)$, while the EAHP produces the weight vector $\mathbf{W}^B = (w_1^B, w_2^B, \ldots, w_n^B)$. The optimized composite weighting vector is obtained as $\mathbf{W} = (w_1, w_2, \ldots, w_n)$. Following the mathematical formalism of the Kullback-Leibler divergence, discriminant information between the new weights and original weights is formulated as:

$$\sum_{j=1}^{n} w_j \ln \frac{w_j}{w_j^A} + \sum_{j=1}^{n} w_j \ln \frac{w_j}{w_j^B} \tag{1}$$

Constructing constraints, we obtain the final optimization model:

$$\min_{j\in\{1,2,\ldots,n\}} F = \sum_{j=1}^{n} w_j \ln \frac{w_j}{w_j^A} + \sum_{j=1}^{n} w_j \ln \frac{w_j}{w_j^B} \tag{2}$$

$$\begin{cases} \sum_{j=1}^{n} w_j = 1 \\ w_j > 0 \; \forall \; j \in 1, 2, \ldots, n \end{cases} \tag{3}$$

For the optimization model above, given the tractable constraints and the continuously differentiable nature of the simplified nonlinear objective function over the feasible region, we employ the method of Lagrange multiplier to derive the composite weight values with analytical precision. The objective function is first simplified as follows:

$$F = \sum_{j=1}^{n} w_j \ln \frac{w_j}{w_j^A} + \sum_{j=1}^{n} w_j \ln \frac{w_j}{w_j^B} = \sum_{j=1}^{n} w_j (2 \ln w_j - (\ln w_j^A w_j^B)) \tag{4}$$

Given the continuity and differentiability of the function across the feasible domain, partial derivatives are derived for each variable. By leveraging the permutation symmetry inherent in w_j, the partial derivative with respect to the j-th weight w_j is computed as a representative case:

$$\begin{cases} \frac{\partial F}{\partial w_j} = 2\ln w_j - \ln z_j + 2 + \lambda \\ \frac{\partial F}{\partial \lambda} = \sum_{j=1}^{n} w_j - 1 \end{cases} \tag{5}$$

For the above system of $n+1$ equations, set the right-hand side of each equation to 0, yielding the following transformation:

$$\frac{\partial F}{\partial w_j} = 0 \Rightarrow \ln w_j - \ln \sqrt{z_j} + 1 + \frac{\lambda}{2} = 0 \Rightarrow \ln w_j = \ln \sqrt{z_j} - \left(1 + \frac{\lambda}{2}\right) \tag{6}$$

Taking the exponential of its both sides, using $\frac{\partial F}{\partial \lambda} = 0$, we obtain equation$\sum_{j=1}^{n} w_j = 1$. Combining this equation with 6, we get:

$$\sum_{j=1}^{n} w_j = e^{-\left(1+\frac{\lambda}{2}\right)} \sum_{j=1}^{n} \sqrt{z_j} \quad \Rightarrow \quad 2 + \lambda = 2\ln(\sum_{j=1}^{n} \sqrt{z_j}) \tag{7}$$

Substituting the expression for 6 and substituting $z_j = w_j^A w_j^B$, the composite weight value is obtained as:

$$w_j = \frac{\sqrt{w_j^A w_j^B}}{\sum_{j=1}^{n} \sqrt{w_j^A w_j^B}}$$

Herein, still using the USA as an example, the weights for CRITIC and EAHP are obtained through solving. To calculate the comprehensive weights, we substituted the previously obtained weights into the formula above, and got the following results:

Table 5: Comprehensive Weights

	S_Gold	L_Gold	S_Total	L_Total	Rate	Events
CRITIC	0.1318	0.1076	0.1073	0.1013	0.3799	0.1719
EAHP	0.1711	0.1320	0.0808	0.052	0.431	0.0328
KL-CEAHP	0.1642	0.1303	0.1017	0.0793	0.4422	0.0821

Finally, in order to simultaneously incorporate the maximization of group utility, minimization of individual regret, and incorporate the decision-makers' subjective preferences, we applied the VIKOR method to evaluate the importance of sports. The basic concept of the VIKOR method is to determine the Positive Ideal Solution (PIS) and Negative Ideal Solution (NIS), then compare the evaluation values of alternative schemes based on their distance from the ideal indicators, selecting the optimal solution. Compared with TOPSIS(Technique for Order of Preference by Similarity to Ideal Solution) and TODIM(an acronym in Portuguese for Interactive and Multicriteria Decision Making), VIKOR offers higher ranking stability and reliability. The pseudocode is as follows:

Algorithm 1 VlseKriterijumska Optimizacija I Kompromisno Resenje (VIKOR)

Input: Weight of each variable w and Evaluation index data b_{ij}
Output: Scheme pros and cons ranking
Step 1: Standardized decision matrix
Step 2: Determine group utility S and individual regret R

$$S_i = \sum_{j=1}^{n} w_j \cdot \frac{\max_i\{b_{ij}\} - b_{ij}}{\max_i\{b_{ij}\} - \min_i\{b_{ij}\}}$$

$$R_i = \max_j \left\{ w_j \cdot \frac{\max_i\{b_{ij}\} - b_{ij}}{\max_i\{b_{ij}\} - \min_i\{b_{ij}\}} \right\}$$

Step 3: Calculate the compromise index value Q ($v \in [0,1]$, the decision coefficient)

$$Q_i = \frac{v(S_i - \min_i\{S_i\})}{\max_i\{S_i\} - \min_i\{S_i\}} + \frac{(1-v)(R_i - \min_i\{R_i\})}{\max_i\{R_i\} - \min_i\{R_i\}}$$

Return: S, R, Q and Scheme ordering

The VIKOR method was employed to calculate the importance ranking of sports. The top six most important sports for the USA are presented in Figure 9:

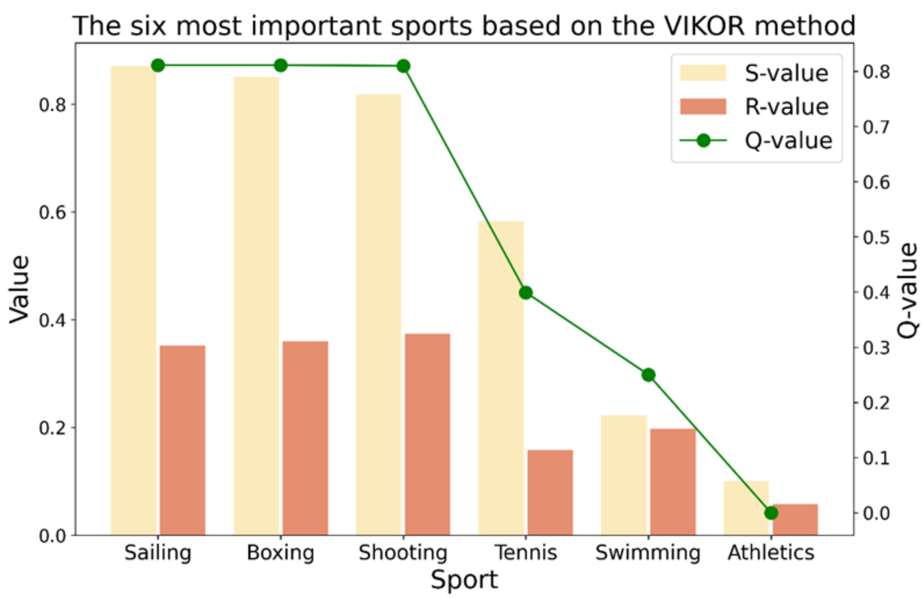


Figure 9: Top Important Sports for the USA

Based on the group utility value S and individual regret value R, the decision index Q was calculated, where a lower Q value indicates a more favorable sport. The figure shows that Athletics is the USA's most important sport, consistent with our prior analysis. Although Shooting has a higher Q value, indicating it is less optimal, it remains among the top sports, reflecting the USA's strong focus and competitive advantage.

For Swimming, the USA has been a dominant force since the 1896 Olympics, accumulating 609 medals in this event, far surpassing other competitors. Tennis follows the same trend, highly valued as well.

4. **How do the events chosen by the home country impact results?**

The "host country" effect is confirmed by both Figure 3 and the significant coefficients of the host country variables in the regression model, demonstrating its substantial impact on medal outcomes.

Data from the 2024 Paris Olympics and 2020 Tokyo Olympics, combined with the KL-CEAHP-VIKOR method's ranking of host country sports, reveal that host countries tend to promote their advantageous (or key) events in the Olympics, thereby increasing the number of events in their strong sports to showcase their unique sports culture.

Thus, host countries prioritize their competitive events, resulting in more gold and total medals, improving their overall performance.

6 Task 2: Exploring the "Great Coach" Effect

6.1 Determining Research Objects

1. Chinese coach *Lang Ping* became the head coach of the U.S. women's volleyball team in 2005, leading the team to a silver medal at the 2008 Beijing Olympics.
2. Australian coach *Alyson Annan* led the Australian women's field hockey team to two Olympic gold medals (1996 Atlanta and 2000 Sydney). She then coached various age-group teams for the Netherlands, becoming the head coach of the Netherlands women's national field hockey team in 2015.
3. French coach *Benoît Vêtu* became the coach of the Chinese cycling team on October 15, 2013. He led the Chinese cycling team to win China's first Olympic gold medal in cycling at the 2016 Rio Olympics.

6.2 Demonstrating the "Great Coach" Effect

To validate the "great coach" effect, we controlled for non-dominant factors by selecting foreign coaches' sports and an appropriate time span, allowing us to compare medal outcomes before and after their involvement.

We first filtered the relevant data and present it in Table 6 below, where the horizontal axis represents the years of the Olympic Games, and the vertical axis shows countries and their medal counts, with the first letter representing each type of medal (e.g., G for gold).

Table 6: Medal Tally for the Specified Events

		1988	1992	1996	2000	2004	2008	2012	2016	2020	2024
USA	G	0	0	0	0	0	0	0	0	1	0
	S	0	0	0	0	0	1	1	0	0	1
	B	0	1	0	0	0	0	0	1	0	0
CHN	G	0	0	0	0	0	0	0	1	2	1
	S	0	0	0	0	1	1	2	0	0	0
	B	0	0	0	1	0	0	1	0	0	0
NED	G	0	0	0	0	0	1	1	0	1	1
	S	0	0	0	0	2	0	0	1	0	0
	B	1	0	1	1	0	0	0	0	0	0

To investigate the existence of the "great coach" effect, we analyze the changes in medal outcomes of the country's coached sports, using the foreign coach's appointment as the time marker. Considering the rules of the Olympic medal ranking system, where the total medal tally is ranked by gold medals first, followed by silver if golds are equal, and then bronze, we assign scores to medals as follows: gold = 25, silver = 10 and bronze = 5. This helps to clearly distinguish the importance of each medal and provides a score for each country in every Olympic edition.

Subsequently, with the score sequence for each country, we apply the Moving T-test (MTT) to examine whether there is a significant effect on the country's medal counts. The mathematical procedure is outlined as follows:

For a score sequence x with n observations, a certain time point is chosen as the reference point based on prior information. The sample sizes of the two sub-sequences, x_1 and x_2, before and after the reference point, are n_1 and n_2 respectively. The means of these sub-sequences are $\bar{x_1}$ and $\bar{x_2}$, and their variances are $\bar{s_1}$ and $\bar{s_1}$.

Define hypotheses as follows:

$$\begin{cases} H_0 : \bar{x_1} = \bar{x_2} \\ H_1 : \bar{x_1} \neq \bar{x_2} \end{cases} \tag{8}$$

Let the significance level α be set to 0.1. Therefore, the test statistic is defined as:

$$t = \frac{\bar{x_1} - \bar{x_2}}{s\sqrt{\frac{1}{n_1} + \frac{1}{n_2}}}, \tag{9}$$

where $s = \sqrt{\frac{n_1 s_1^2 + n_2 s_2^2}{n_1 + n_2 - 2}}$, and the equation 9 follows a t-distribution with degrees of freedom $v = n_1 + n_2 - 2$.

Taking the USA as an example, its score sequence from 1988 to 2024 Olympics is $x = [0, 5, 0, 0, 0, 10, 10, 5, 25, 10]$. Considering that Lang Ping became the head coach of the U.S women's volleyball team in 2005, we specify 2005 as the reference point to divide the sequence into two sub-sequences: $x_1 = [0, 5, 0, 0, 0]$ and $x_2 = [10, 10, 5, 25, 10]$. Their means are $\bar{x_1} = 1$ and $\bar{x_2} = 12.5$, the variances are $s_1^2 = 5$ and $s_2^2 = 57.5$, and the pooled variance is $s = 6.25$. We can then calculate the t value:

$$t = 3.1112$$

, according to equation 9.

Its corresponding P value is 0.0144, which is smaller than 0.05. From this, we have sufficient evidence to reject the null hypothesis (8) of equal means between the two sub-sequences, indicating a significant difference in the U.S. women's volleyball team's medal outcomes before and after Lang Ping became the head coach, and indirectly supporting the existence of the "great coach" effect.

Similarly, we repeated the Moving T-test on CHN and NED. The sub-sequences of CHN are $[0, 0, 0, 1, 10, 10, 21]$ and $[25, 50, 25]$. And those of NED are $[1, 0, 1, 1, 20, 25]$ and $[25, 10, 25, 25]$. Their calculated P values are 0.0043 and 0.0761, both smaller than 0.1. Again, we demonstrated the the existence of the "great coach" effect.

6.3 Illustrating the Specificity of the "Great Coach" Effect

In the previous section, we demonstrated the existence of the "great coach" effect, showing that the decision to hire foreign coaches may lead to significant improvements in the sports they oversee. However, to quantify its contribution to national-level medal outcomes, it is pivotal to first investigate whether the "great coach" effect also influences sports managed by non-foreign coaches. This would help eliminate potential interaction effects within the features when building subsequent quantitative models.

To begin with, we follow a similar approach by selecting relevant data. For better consistence, we focus on the three different sports corresponding to the three countries discussed in the previous section for our analysis.

Continuing with the USA as an instance, Lang Ping became the head coach of the U.S. women's volleyball team in 2005, with no notable foreign coaches joining the basketball or cycling teams during the same period. Therefore, we can investigate whether the performance of the basketball or cycling teams showed any significant changes during this time span. If such changes exist, then we will explore whether the "great coach" effect accounted for these variations.

Essentially, this analysis seeks to determine whether the "great coach" effect could also explain changes in sports managed by non-foreign coaches, thus exploring the causal relationship. Since the U.S. women's volleyball team's performance sequence in the previous section already reflects the impact of the "great coach" effect, we can similarly construct performance sequences for the basketball and cycling sports. We will then use the Granger causality test to examine whether the "great coach" effect influences the competitiveness of these other sports. The mathematical procedure is described outlined below:

For two sequence data x_t and y_t, assuming that x_t is an independent variable and y_t is the dependent variable, we can express their relationship this way:

$$y_t = a_0 + a_1 x_{t-1} + a_2 y_{t-1} + u_{1t}$$

Define hypotheses:

$$\begin{cases} H_0 : \alpha_1 = 0 \\ H_1 : \alpha_1 \neq 0 \end{cases} \tag{10}$$

Rejecting the null hypothesis H_0 indicates that there is sufficient evidence to suggest x_t as the cause of y_t, since when controlling for the lagged values of y_t, if the lagged values of x_t still contribute to explaining the current variation in y_t, then x_t is considered to have a causal effect on y_t.

Returning to the focus of this section, we can treat the performance sequences of *the U.S. basketball or cycling teams* as the dependent variable y_t, with *the U.S. women's volleyball team* performance sequence as the independent variable x_t. By fitting the Granger causality test expression, we can assess the significance of the corresponding coefficients.

The results show that the Granger causality test P value for the U.S. basketball team is 0.178, and for the cycling team, it is 0.327. Since both P values are greater than 0.1, we cannot reject the null hypothesis. This suggests that changes in the performance of the U.S. women's volleyball team are not caused by changes in the performance of the basketball or cycling teams. Finally, it indirectly supports the idea that the "great coach" effect does not influence other sports managed by non-foreign coaches.

6.4 Quantifying the "Great Coach" Effect

Although the previous section showed that the "Great Coach" effect does not impact the competitiveness of other sports, a country's overall medal count is influenced by multiple sports, each with various contributing factors. To quantify the impact of the "great coach" effect, we model its influence on foreign-coached sports.

Continuing with the USA, we examine the impact of the "great coach" effect on the U.S. Women's Volleyball team from 1988 to 2024. The total medal count is the dependent variable, with the "Great Coach" effect as an intervention variable. We construct an Interrupted Time Series Analysis (ITSA) model to evaluate the pre- and post-intervention regression changes.

The ITSA model is as follows:

$$Y_t = 0.324 - 0.033 \cdot T_t + 0.519 \cdot X_t + 0.05 \cdot X_t T_t,$$

where T_t is the time variable, X_t is the intervention dummy variable, and $X_t T_t$ represents their interaction effect. Their coefficients are significant at the 0.1 level, with the model's $R^2 = 0.714$, confirming a positive impact of the "great coach" effect.

Similarly, for China's cycling and Netherlands' women's hockey, which had prior medal records but lacked gold medals, the model shows a significant positive impact of the "great coach" effect on increasing gold medals.

Thus, the "great coach" effect notably resulted in the USA Women's Volleyball team's medal count increase, especially the breakthrough from 0 to 1 gold medal. For China's cycling and the Netherlands' women's hockey, it boosted gold medal counts from 0 to 1.

Countries without prior medal achievements, or with no gold medals, should consider investing in foreign coaches for higher medal probabilities and quality.

7 Task 3: Additional Insights into Olympic Medal Counts

- **Long-Term Historical Medal Counts Do Not Provide Expected Contribution for Prediction:** In the feature engineering stage, we clarified the contribution of the current athletes' information to the prediction model. However, the information regarding the participating athletes was incomplete. Thus, historical medal data, which reflects the strength and trend of a country's performance in sports, is also pivotal. In our experiments, short-term historical medal counts indeed contributed significantly to the prediction model, while long-term historical medal counts may restrict the country's breakthrough in specific fields. To improve this situation, additional data sources need to be introduced to further optimize feature engineering.

- **A Country's Gender Distribution of Athletes Does Not Significantly Affect Medal Counts on The Whole:** In 4, we used the Wilcoxon rank-sum test to assess whether gender distribution in some countries' medal counts had a significant impact on medal tally. For example, in France, among the non-medalists, gender distribution had a significant effect at $\alpha = 0.01$. However, in the prediction models we constructed, the gender distribution feature was not significant for most countries. Hence, gender distribution is not a crucial factor in overall prediction.
- **Type of Events Has a Significant Impact on Medal Distribution:** As observed in the analysis of the importance of sports in 5.3, our analysis indicates that different sports contribute significantly differently to medal counts, and the dominant sports vary across countries. Additionally, in conjunction with the "host country" feature, host nations tend to prioritize events in which they excel to increase their competitiveness. This factor should be given special consideration in prediction models.
- **The Impact of the "Great Coach" Effect Is Significant:** Our experimental results suggested that the "great coach" effect can lead to a breakthrough in the total medal count from 0 to 1. It may also elevate traditional strong sports, resulting in a significant increase in gold medals from 0 to 1. This effect could provide breakthrough opportunities for countries that have not yet won Olympic medals.
- **Recommendations for National Olympic Committees Based on Model Predictions:**
  - **Introducing Renowned Coaches:** As shown in the results of the second question, the "Great Coach" effect helps achieve breakthroughs in the total medal count from 0 to 1 and significant increases in gold medals. NOCs should consider bringing in internationally renowned coaches to help develop the best strategies for their athletes.
  - **Investment Strategy for Strong Sports:** NOCs should invest deeply in sports where the country already has a competitive advantage, which can increase the likelihood of winning higher-tier medals (gold).
  - **Investment Strategy for Potential Sports:** NOCs should increase resource allocation to sports with high potential and good recent performance to expand the total medal count.
  - **Comprehensive Long- and Short-Term Consideration:** Long-term historical medal counts reflect the country's traditional strengths. However, with the evolving landscape of Olympic competition and team changes, it is necessary to focus on short-term adjustments to align with the current trends of the Olympics.

8 Sensitivity Analysis

The evaluation of important sports plays a crucial role in various aspects, such as predicting the next medal count of a country, formulating the current National Olympic Committee's training strategies, and determining investment strategies for dominant and potential sports, the stability of the important sport evaluation model (KL-CEAHP-VIKOR) constructed in the first question is of the essence. Thus, it is necessary to perform sensitivity tuning.

In the weighting process, KL-CEAHP combines mixed objective and subjective weighting, aiming to minimize the combined discriminant information. This results in the minimization of the risk of misclassification, with its strong stability. In the VIKOR method, to account for the maximization of group utility, the minimization of individual regret and decision makers' subjective preferences, a decision variable v is used to denote subjective preferences. In 5.3, a compromise evaluation was conducted using $v = 0.5$. However, considering the subjective preferences for sports in real scenarios, it is essential to adjust v to examine the model's stability.

In this analysis, the values of v were set to $\{0.35, 0.4, 0.45, 0.5, 0.55, 0.6, 0.65\}$ as parameter fluctuations. Taking the USA as an example, we conducted a sensitivity analysis on the group utility value S, individual regret value R, and decision index Q for the top two important sports. The results of this analysis are visualized as follows:

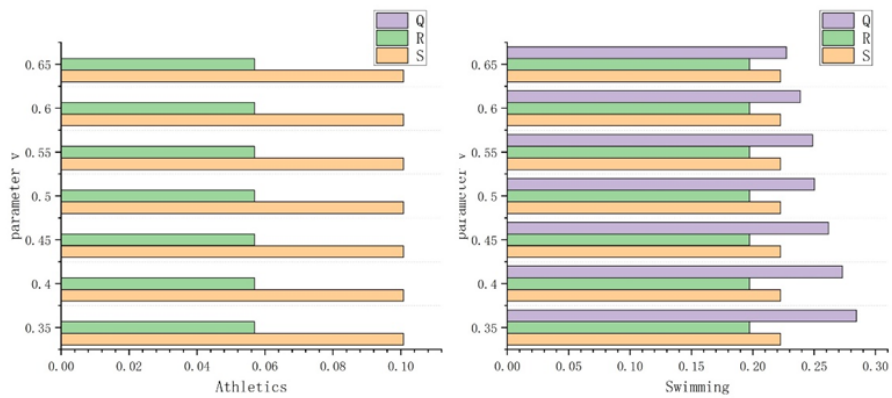


Figure 10: Sensitivity Analysis

From the graphical results above, it can be observed that Athletics and Swimming have consistently remained the two most important sports for the USA. In Athletics, regardless of v fluctuations, the group utility value S, individual regret value R, and decision index Q remain stable, indicating strong stability. For Swimming, as v increases, the preference for the group utility value S grows, while the decision index Q continuously decreases, fluctuations remaining within a 10% range. This indicates that Swimming holds significant importance in the overall sports events of the USA, and its evaluation is stable.

In conclusion, the experiments demonstrate that our evaluation model retains high stability even with large fluctuations in subjective preferences, providing reliable support for the Olympic Committee.

9 Conclusion

Comprehensively, following conventional procedures, we integrated multiple innovative and original approaches to, first, break down problems and then tackle them one by one. Eventually, we achieved expected and satisfying outcomes.

9.1 Strengths

1. **Model Adaptability and Stability:** Based on the model's prerequisites, we selected the most appropriate model for the problem and perform sensitivity analysis to demonstrate that its results are stable and reliable.
2. **Model Rationality and Innovation:** From both objective and subjective perspectives, we comprehensively consider the nature of the problem and the strengths and weaknesses of various baseline models. By leveraging the advantages of each, we innovatively construct a new model better suited for solving practical problems, achieving results that align with real-world conditions.

9.2 Weaknesses

Insufficient Data: To predict medal counts more accurately, factors such as current news and social events should also be considered, incorporating multimodal input information to adjust the prediction trends.

References

[1] Allen, Sian V, and Will G Hopkins. "Age of Peak Competitive Performance of Elite Athletes: A Systematic Review." Sports medicine (Auckland, N.Z.) vol. 45,10 (2015): 1431-41. doi:10.1007/s40279-015-0354-3

[2] Forrest, D., Sanz, I., & Tena, J.D. (2010). Forecasting national team medal totals at the Summer Olympic Games. *International Journal of Forecasting*, 26(3), 576-588. `https://www.sciencedirect.com/science/article/pii/S0169207009002088`. Accessed: Jan. 27, 2025. `https://doi.org/10.1016/j.ijforecast.2009.12.007`.

[3] Schlembach, Christoph & Schmidt, Sascha & Schreyer, Dominik & Wunderlich, Linus. (2020). Forecasting the Olympic Medal Distribution during a Pandemic: A Socio-Economic Machine Learning Model. SSRN Electronic Journal. 10.2139/ssrn.3745595.

[4] Nielsen, (2024). Virtual Medal Table Forecast. Nielsen. `https://www.nielsen.com/news-center/2024/virtual-medal-table-forecast/`. Accessed: Jan. 27, 2025.

Report on Use of AI

1. DeepSeek DeepSeek-V3 (Dec 2023 version, DeepSeek-V3)

   Query: What should I be careful of when writing an MCM paper? Give me a brief illustration within one paragraph.

   Output: When writing an MCM (Mathematical Contest in Modeling) paper, be careful to clearly define the problem, state your assumptions, and justify your approach. Use precise mathematical language and ensure your models are well-developed and validated. Present results logically with supporting data, graphs, or tables, and discuss their implications. Avoid overly complex jargon that may obscure understanding. Proofread for clarity, coherence, and adherence to the contest's formatting guidelines. Finally, cite all sources properly and ensure your work is original and collaborative.