Exploration of Olympic Medal Distribution: A Combined Application of BP Neural Network and Statistical Methods

Ziyu Gao, Yonggang Guo, Zhifeng Chen, Ziyan Cui, Qizheng Zhang
Shijiazhuang Tiedao University, Shijiazhuang, Hebei, China

Abstract: This paper examines Olympic medal distribution using various models and methods to provide decision-making insights for the International Olympic Committee and national sports bodies. In Question 1, a BP neural network model predicts 2028 medal counts and changes per country, incorporating factors like athlete participation, gender ratio, home-field advantage, and sports types, while estimating first-win probabilities for "0 prize" countries after addressing missing values and outliers. Correlation analysis then links sports to medals, e.g., Japan's edge in judo and wrestling. For Question 2, the BP neural network and mathematical statistics identify "great coach" effects in countries like Japan and Mexico, with sensitivity analysis pinpointing investment-worthy sports, such as judo and gymnastics in Japan. Question 3 summarizes medal distribution patterns from prior models, offering the IOC resource allocation and strategy recommendations. The study's multi-factor model ensures high accuracy, identifies key influences, and suits all country types. Future dynamic modeling with real-time data could boost robustness. Results optimize sports resource allocation and inform data-driven decisions in finance, healthcare, and beyond.

Keywords: Prediction and Optimization of Olympic Events; BP Neural Network Models; Correlation Analysis; The "Great Coach" Effect; Mathematical Statistics

1. Introduction

1.1 Problem Background

The Olympics, the top global sports event, blend athletic peaks with historical and cultural depth. Medals embody excellence and the motto "higher, faster, stronger," symbolizing progress and cooperation. Medal distribution reflects national sports strength, competitiveness, and global shifts. Analyzing it reveals capabilities, trends, and cultural exchange, while driving innovation and sports science [1]. Studying medals offers practical and academic worth.

1.2 Restatement of the Problem

For Question 1: Build a model to predict the 2028 Olympic medal table, check accuracy, and track medal count changes per country. Analyze non-medal countries for first-gold odds and link program factors to medals [1], guiding Olympic Committees on event picks. For Question 2: Develop a model to assess the "Great Coach" effect, pick three countries, find best programs for coaching investment, and evaluate medal changes [2]. For Question 3: Summarize insights from these models on medal counts and their value to the IOC.

1.3 Analysis of the problems

Analysis of Problem 1

This problem has three sub-problems: (1) predicting 2028 Los Angeles Games medal counts and changes per country; (2) evaluating first-win odds for non-medal countries; (3) analyzing event types' correlation with medal wins. Annex data is preprocessed for accuracy by fixing missing values and outliers with official data. Key indicators-athlete participation, event types, home advantage-are mined, and a BP neural network quantifies their medal impact, refined by error corrections. The model forecasts 2028 medals, tracks changes from past data, and estimates "0" medal countries' win chances. Correlation analysis uses these indicators as variables to pinpoint major medal influences, aiding strategic planning, as shown in Figure 1.

Analysis of Problem 2

Question 2 has two parts: (1) spotting "great coach" signs, (2) picking three countries for coaching investment. Olympic data shows coaching differs by country. BP neural network predicts medals sans coaching; actual-predicted gaps show errors and coaching impact, with big deviations hinting its role [2]. We compute medals, check deviations, use 3σ for outliers signaling coaching effects. For three countries, we assess event medals, sensitivity, and contributions, comparing gains to plan ahead, as shown in Figure 2.

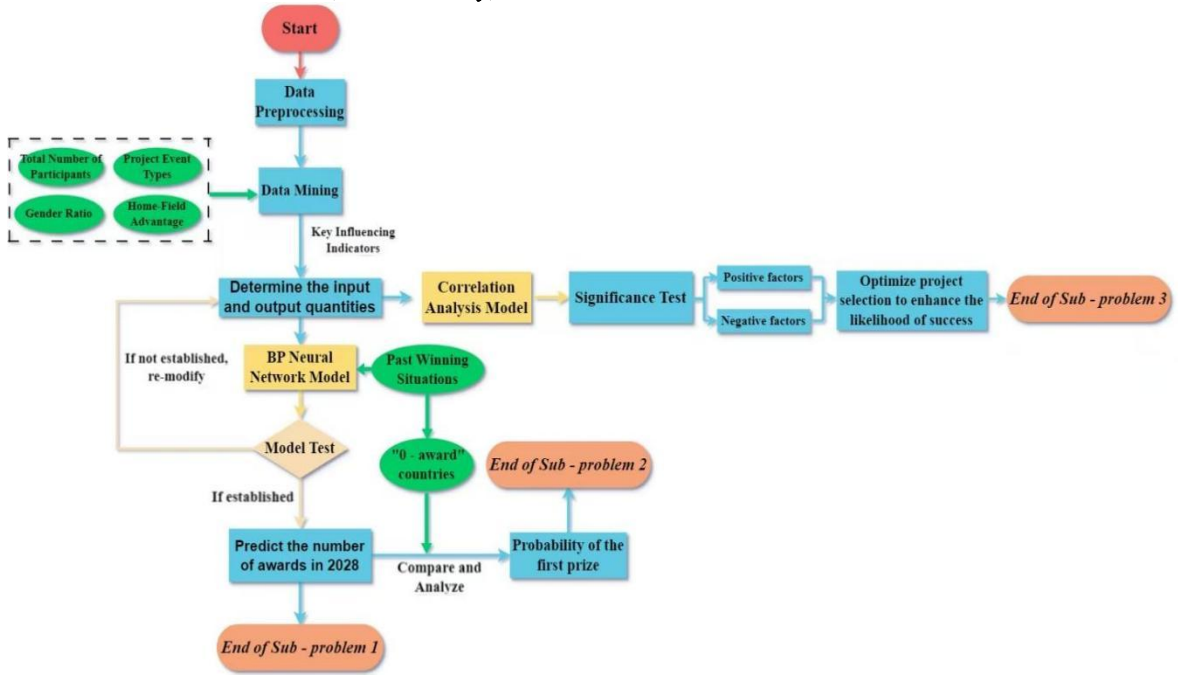

Figure 1. Road Map for Question 1 Research

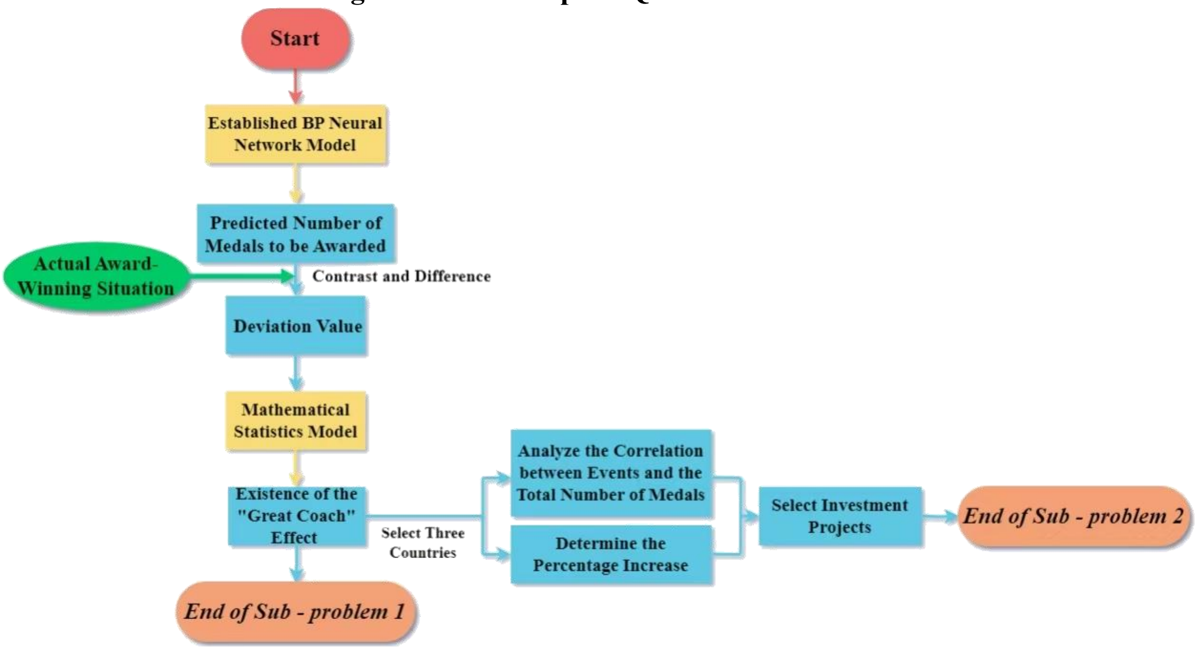

Figure 2. Road Map for Question 2 Research

Analysis of Problem 3

Question 3 sums up medal count insights from Questions 1 and 2's models, showing their IOC value. Question 1 ties 2028 medal trends to participation, investment, and home advantage. Question 2 links the "Great Coach" effect to medal gains in key sports, boosting results. These help IOC track shifts, plan events [3]. Medal model spots 2028 gainers or first-timers, focusing resources on emerging nations. Coaching analysis flags sports lifted by top coaches, shaping allocation. Models predict trends, support IOC strategy.

2. Model Assumptions and Notation

2.1 Model Assumptions

The model simplifies complexity with these assumptions:

(1) Appendix-extracted factors cover all key

variables significantly tied to medal counts.
(2) External factors like refereeing, athlete selection, and weather have negligible impact on predictions.
(3) Past medal patterns are regular and offer a predictive reference.
(4) Processed data fully and accurately reflect each country's Olympic situation.

2.1.1 Description of symbols

The conformity used in this paper is listed in Table 1.

Table 1. Notations Used in This Paper

Symbol	Description
x_{ij}	total number of participants from country i in the j th Olympics
x_k	number of participants in the kth competition event
y_{jk}	total number of entries in the kth competition event of the jth Games
m	total number of competition events
n	total number of Olympic Games held
z_{ij}	home field advantage of the ith country in the jth Olympic Games
p_{ij}	Percentage of participation by gender
G_{ij}, S_{ij}, B_{ij}	indicates the number of gold, silver, and bronze medals won
T_{ij}	total number of medals won by the ith country at the jth Games
W	relative error
p	probability of winning a medal
r	Pearson correlation coefficient (statistic)
Y	actual awards for the ith country at the jth Olympiad
$\hat{Y}$	predicted awards for country i in the jth Olympiad
δ	relative deviation between actual and predicted values
d	increase generated by coaching benefits in the kth competition event

2.1.2 Data pre-processing

In order to ensure the accuracy and reliability of the prediction results, it is first necessary to carry out comprehensive data processing of the data provided in the annex. In the process of data preprocessing, the main focus is on two types of data that need to be corrected, namely, missing values and outliers.

2.1.3 Handling of missing values

Missing values arise from unrecorded data due to omissions or poor collection. Excel scans data to spot blanks and inconsistencies in initial processing. We use official tournament sites and databases for accurate data, cross-checking to fill gaps, ensuring integrity, minimizing bias, and boosting prediction accuracy and reliability.

2.1.4 Handling of outliers

Outliers, far from data norms, stem from errors, extreme events, or unique factors, risking model accuracy like missing values. Detecting and fixing them boosts performance. Logical anomaly rules spot errors in attachments, ensuring integrity. Suspected outliers are checked and corrected with sources. Steps: (1) Match geography, fixing region-country errors via IOC data (e.g., Soviet Union). (2) Verify athletes' event records over years. (3) Align category medal totals with overall, correcting misreports.

2.2 Modeling and Solution

Problem 1 modeling and solution

Medal count prediction and change modeling

(1) Definition of input and output variables

After the above data preprocessing, we extracted four influencing factors through in-depth analysis and data mining of the corrected data: the total number of athletes competing in each country, the type of program events, the gender share, and home field advantage or not. This study concludes that the above factors are key indicators that influence the number of medals won by each country each year, as defined below:

A. Total number of participants

In Olympic events, the degree of importance and commitment of countries in sports programs and competitive ability can often be visualized through the number of participants. Therefore, analyzing the number of participants in individual events can not only reflect the allocation of sports resources by each country, but also indirectly reveal the potential of each country to win medals in the events. Define the total number of participants of country i in the j th Olympic Games x_{ij} x_{ij} .

$$x_{ij} = \sum_{k=1}^{m} x_k \tag{1}$$

Where x_k denotes the number of participants of the corresponding country in the kth event; m denotes the total number of events in the competition.

B. Number of Program Events

In Olympic sports, the total number of events gauges competition intensity and country investment. It reflects participation and preparation levels, with more events signaling fiercer competition. As competition rises, countries invest more, boosting their medal chances.

To further quantify this influence, let the total number of k th events at the j th Games be y_{ij}. In addition, this factor indicator should be considered separately for gold, silver, and bronze medals, which are:

$$T_{ij} = G_{ij} + S_{ij} + B_{ij} \tag{2}$$

where T_{ij} denotes the total number of medals won by the ith country at the jth Games, and C_{ij}, S_{ij}, and B_{ij} denote the number of gold, silver, and bronze medals, respectively, won by the ith country at the jth Games., and the number of bronze medals.

C. Home Field Advantage

Home field advantage usually refers to the additional advantage that a participating team may gain due to geography, psychology and other factors when hosting a competition in its own country or region. In order to quantify this advantage, this paper indicates the existence of home field advantage by introducing a 0-1 variable defined as follows:

Let Z_{ij} be a binary 0-1 variable indicating the home field situation of the ith country in the jth Games. If the host country of the jth Games happens to be the ith country, then its venue is also in the corresponding country thereby positively affecting the country's winning performance, it can be considered as having a home field advantage and Z_{ij} = 1. Conversely, it does not have a home field and Z_{ij} = 0.

The definition of this 0-1 variable allows home field advantage to be quantified by mathematical modeling and taken into account as a factor in further modeling, leading to in-depth analyses of Olympic Games in terms of gold medal predictions, team performance, etc.

D. Gender Share

Gender share significantly influences Olympic medal distribution, particularly in sports where male-female performance gaps affect outcomes. Changes in the gender [3] ratio can directly alter medal trends, especially in sports with notable performance disparities. Accounting for gender share improves the accuracy of predicting a country's Olympic performance. Define the gender share P_{ij} as the ratio of male to female athletes for the ith country in the jth Olympics. Based on four key factors, predictive indicators (*x, y, z, p*) are derived.

(2) Establishment of BP neural network model

Establishing the above key influence index data as input variables, selecting three types of Olympic gold, silver and bronze as output variables, using BP neural network model to quantify the mapping relationship between each input variable and the output variables [4], so as to construct a prediction model about the number of Olympic medals, the specific operation steps are as follows:

Step1: Define the connection right between the input layer and the intermediate layer w_1 ,The connection right from the intermediate layer to the output layer w_2 ,set the (-1,1) interval to provide random values, assign random values to the output queue values of each unit within the neural network, set the error function e and the maximum number of training times M. Randomly select k network input vectors $L_k = \left(l_1^k, l_2^k, \ldots, l_n^k\right)$ and corresponding network output vectors $A_k = \left(a_1^k, a_2^k, \ldots, a_n^k\right)$ which are provided to the neural network.

Step2: Obtain the vector of input middle layer units s_j through the output samples, connection weights and thresholds, and calculate the transfer output b_j.

$$b_j = f\left(s_j\right), j = 1,2,\ldots,p \tag{3}$$

$$s_j = \sum_{i=1}^{n} w_{ij} a_i - \theta_j, j = 1,2,\ldots,p \tag{4}$$

Step3: Calculate the output vector $T_t = \left(y_1, y_2, y_3, \ldots, y_t\right)$ through the transfer function and network expectation, define the actual output value as c_t, unit error d_t^k.

$$d_t^k = \left(y_t^k - C_t\right) \cdot C_t\left(1 - C_t\right), t = 1,2,\ldots,p \tag{5}$$

Step4: Backtrack the connection process, combine the connection weights, unit error d_t and intermediate layer output b_j to get the intermediate layer generalized error e_t^k, and correct the weights and computational thresholds of the output, intermediate, and input layers, respectively, by the gradient descent method.

$$e_t^k = \left[\sum_{i=1}^{q} d_t v_{jt}\right] b_j \left(1 - b_j\right) \quad (6)$$

Step5:Re-select new learning samples and loop the above steps to train *m* samples. The whole process is usually repeated several times (i.e., multiple epochs) until the error converges.[5]

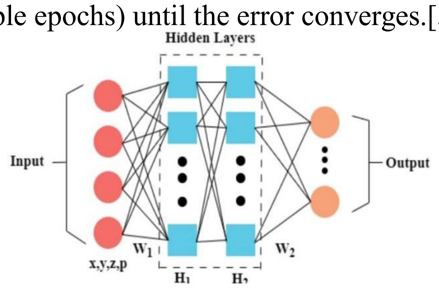

Figure 3. Schematic Diagram of Neural Network Structure

(3) Prediction of input variables

Using the BP neural network framework, this paper predicts key indicators-total participants, events, gender ratio, and home-field advantage-for each country in the 2028 Olympics, as shown in Figure 3. To ensure model accuracy and robustness, the following steps are detailed:

Step 1: Select n historical datasets as training set X and the n+1th as output set Y, allowing the model to learn patterns and predict future data effectively, with varied time nodes enhancing applicability and precision.

Step 2: Design the BP neural network by setting layer numbers and training iterations based on data traits and task needs, adjusting neurons and structure to capture input features fully.

Step 3: Predict output Y′ from input X using the trained network, then compute the relative error between predicted and actual values with the error formula:

$$w_1 = \frac{Y' - Y}{Y} \quad (7)$$

Where, w_1 represents the relative error, Y' represents the result predicted by the neural network model, and Y is the actual observation data.[6] By calculating and analyzing the error through the MATLAB software, the performance of the model can be evaluated and the model can be trained to obtain the predicted values of the key impact parameters affecting the 2028.The 2028 Olympic Games are held in Los Angeles, USA. Therefore, it can be determined that the U.S. team has a home field advantage. The 2028 input variables for the partial predictions obtained are summarized as follows in Figure 4, Figure 5, Figure 6:

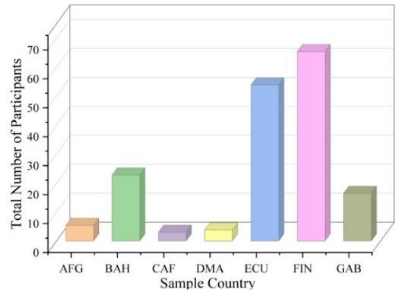

Figure 4. The Total Number of Participants from the Sample Countries

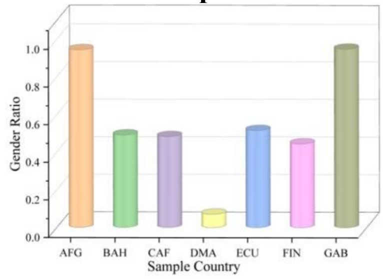

Figure 5. Gender Ratio from the Sample Countries

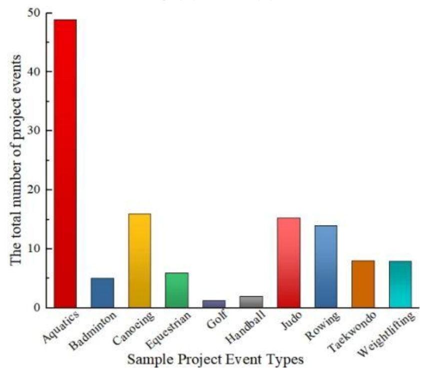

Figure 6. The total Number of the Sample Project Event Types

Through the above steps and methods, the BP neural network can provide a prediction of the performance of each country in the future Olympic Games, providing data support for relevant decisions and helping to further analyze the potential of each country to compete

(4) Prediction and test of output variables

The total number of participants x_{ij}, the total number of events in the program y_{ij}, the home field advantage Z_{ij}, and the gender share pij will be used as input variables, and the number of gold, silver, and bronze medals C_{ij}, S_{ij}, *and* B_{ij} of each country in the Olympic Games will be used as output variables. In this way, this paper will try to build a more accurate prediction model for simulating and analyzing the performance of each country in the Olympic Games, and use it to optimize the future preparation strategy.

By substituting the data in the above table into the BP neural network model for prediction, the predicted winning data were obtained as C_{ij} , S_{ij} , *and* B_{ij} and then repeated the training, and finally obtained to the relevant parameters and mapping functions.

In order to analyze the accuracy of the model, it is necessary to carry out an error test on the prediction results , defined.

$$w_2 = \frac{1}{3n}\sum_{i=1}^{n} | G_i - G'_i | + | S'_i - S_i | + | B'_i - B_i | \quad (8)$$

Using the above model, the final solution to get the medals for each country in 2028 is as follows:

Table 2. Projected 2028 Awards for Selected Countries

Name of country	Number of medals	Name of country	Number of medals
AFG	0	NCA	25
BAH	13	OMA	0
CAF	0	PAN	4
DMA	2	QAT	0
EOR	0	RUS	36
FIN	26	SCG	0
GEQ	0	TAN	9
HAI	3	UGA	16
IOA	6	VIE	0
JAM	53	WIF	14
KAZ	0	YMD	4
LAO	4	ZAM	0

Then the predicted 2028 awards of each country, the data is compared with the 2024 awards of each country. Thus, the changes in the medals won by each country are obtained, and the specific progress and regression are shown in Figure 7:

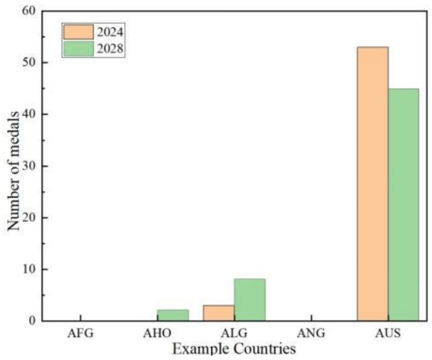

Figure 7. Comparison of Awards in 2024 and 2028 for Selected Countries

From the above figure 7, we can see that the countries with progress are AHO, ALG; the countries with regression are AUS; and the countries with stable level of awards are AFG,ANG, and more data can be found in the appendix.

Modeling the probability of first-time winners

The next step analyzes the probability of first-time medal wins for countries with no prior medals, assessing their likelihood of success in future Olympics. By deeply examining “0 prize” countries’ traits, we estimate their chances of winning a medal. These countries are identified by filtering historical data:

AHO, AIN, AND, ANG, ANT, ANZ, ARU, ASA, BDI, BHU, BIH, BIZ, BLR, BRN, CAF, CGO, CIV, CMR, COD, COK, COM, CPV, CRT, DMA, ESA, ESP, EUN, GAM, GBS, GDR, GEQ, GUI, GUM, HON, ISV, KGZ, KIR, KOR, LAO, LBN, LBR, LCA, LES, LIB, LIE, MAR, MAW, MDA, MHL, MKD, MLI, MNE, MRI, MTN, MYA, NFL, NRU, OMA, PLE, PLW, PNG, PRK, RHO, ROT, RSA, RWA, SCG, SOM, SSD, STP, SUI, SVK, TKM, TLS, TPE, URS, VAN, VNM, YAR, YEM, YMD, ZIM.

Define s_j the number of awards in 2028 for countries that have not won a medal as having:

$$s_j = G_j + S_j + B_j \quad (9)$$

The probability of winning p_j is defined as follows:

$$p_j = \frac{s_j - \min s_j}{\max s_j + \min s_j} \quad (10)$$

Substituting the projected 2028 awards for each country into the above equation, the probability of each "0 prize" country winning a medal for the first time can be obtained. Only countries with a probability of winning a medal higher than 50% are listed here; full information is provided in the Figure 8:

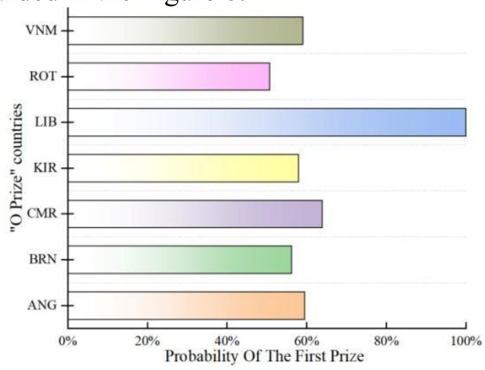

Figure 8. Probability of Receiving the First Prize in Some "0 Prize" Countries

Countries with over 60% win probability are likely to medal in 2028, and over 95% are certain to win. Thus, VNM and CMR are poised for first wins, with LIB highly likely to medal first.

Project and Medal Association Model

To examine the link between Olympic sports and medals won by countries, a correlation analysis model can identify influential sports and assess how sport choices affect medal performance. Steps include:

Step 1: Collect data from past Olympics per country, including four key impact indicator categories (see data summary table above).

Step 2: Calculate linear correlation coefficients between variables using the Pearson correlation coefficient.

$$r = \frac{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\left(y_i - \bar{y}\right)}{\sqrt{\sum_{i=1}^{n}\left(x_i - \bar{x}\right)\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}} \tag{11}$$

Where x_i and y_i represent the correlation data points between the two variables (e.g., the number of medals for a country and the value of its correlation factor); $\bar{x}$ and $\bar{y}$ are the mean values of the respective variables.

The value of correlation coefficient r is (-1,1), the larger the positive value, the stronger the positive correlation; on the contrary [7], the larger the negative value, the stronger the negative correlation; r is closer to 0, the weaker the correlation.

Step3: After calculating the correlation coefficient, it is necessary to test its significance to determine whether the relationship between the two variables is real. The following hypotheses are carried out:

The original hypothesis: $H_0 : r = 0$ assumes that there is no significant linear relationship between the two;

Alternative hypothesis: $H_1 : r \neq 0$ assume that there is a significant linear relationship between the two.

Step4**:** The commonly used statistic for significance test is t-test [8], which is calculated as:

$$t = \frac{r\sqrt{n-2}}{\sqrt{1-r^2}} \tag{12}$$

where r is the Pearson correlation coefficient and n is the sample size.

The t-value is compared to the critical value at n-2 degrees of freedom: if higher, the null hypothesis is rejected, showing correlation; if lower, it holds, indicating no linear link. This spots key programs affecting medals, aiding resource allocation for future Olympics. Japan's correlation model results follow in the Figure 9:

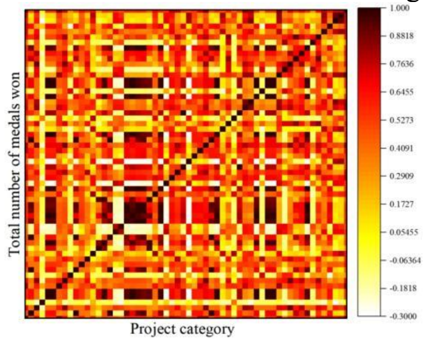

Figure 9. Heat Map of Correlation between Projects and Number of Awards Per Country

Analyzing the chart's numbering with sports names reveals Japan's strong Olympic correlation with Judo, Wrestling, Gymnastics, and Swimming. Japan excels traditionally in judo and wrestling, key strengths, while gymnastics and swimming have recently shone, with athletes earning many medals in major events.

Problem 2 modeling and solution

Modeling the Existence of the "Great Coach" Effect

The coaching effect reflects coaches' impact on athletes' training and competition strategies, improving Olympic performance via training techniques, psychological support, and tactical adjustments. Countries with better coaching systems often see superior athlete results. A BP neural network model predicts medal counts using key indicators, excluding coaching effects, with errors from model predictions and coaching influence. A significant excess of actual medals over predictions signals a strong coaching effect, elevating performance beyond expectations. This effect is quantified by the gap between actual and predicted results, where a large deviation highlights its potency in the data. Define the data of the actual awards of each country in each Olympic Games from the annex as the data Y_{ij}, and the data obtained from the validated and accurate model as the expected data $\hat{Y}_{ij}$, compare the two, and if there is a case of $\hat{Y}_{ij} > Y_{ij}$, it can be regarded as the benefit of the "Great Coach", and define the value of the relative bias arising from the emergence of this effect as δ_{ij}, then there are:[8]

$$\delta_{ij} = \frac{Y_{ij} - \hat{Y}_{ij}}{\hat{Y}_{ij}} \tag{13}$$

Through the principle of 3σ to obtain the relative deviation data δ for mathematical statistics, define each country's prediction deviation data mean is μ, standard deviation is σ, then there are:

$$\mu = \frac{1}{n}\sum_{i=1}^{n}\delta_{ij} \tag{14}$$

$$\sigma = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\delta_{ij} - \mu\right)^2} \tag{15}$$

Where δ_{ij} is the relative deviation value generated by the awards won by the ith country in the jth Games.

Assuming that the error satisfies normal distribution, based on the principle of 3σ, the upper and lower limits of the deviation value are calculated [8], and if the deviation value exceeds this range, it is regarded as an outlier, and the specific formula is as follows:

$$UpperBound = \mu + 3\sigma \tag{16}$$

$$LowerBound = \mu - 3\sigma \tag{17}$$

The "Great Coach" benefit can be regarded as an enhancement to the performance in general, and therefore only focus on the results exceeding the upper limit: AUS, FRA, USA, RUS, GBR, the specific occurrence and year comparison is as shown in Table 3:

Table 3. Years in which the "Great Coach" Effect Occurred

Year	Name of country
1900	SWE, USA
1904	GBR
1908	USA
1912	USA
1920	GER
1928	GBR, USA
1936	GBR, USA
1948	GER
1960	RUS
1964	GBR, UAE
1980	USA
2000	SEN
2012	USA

The "Great Coach" Program Investment Model

The coaching effect, akin to team size growth, lifts performance via better skills or training. We assume a fixed effect per program, adding d participants per sport for a "great coach." The prediction model tests this on 2024 medal counts for three countries. Sensitivity analysis finds coaching-responsive events with big medal gains. Quantitative analysis shows program contributions, spotting coaching-sensitive ones. Correlation analysis links medals to factors, using coefficients to flag high-impact programs. From Question 1, coaching-susceptible, medal-boosting items are picked. Japan, Mexico, and Japan are selected for "Great Coach" investment guidance, as shown in Figure 10:

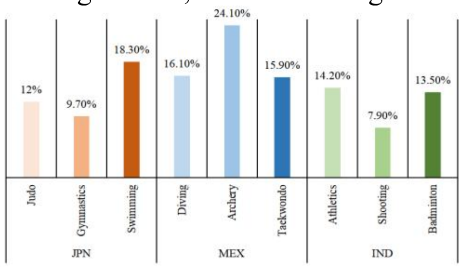

Figure 10. Heat Coach Effect Investment Program in Selected Countries

The chart shows that Japan excels in Judo and Swimming, while Mexico stands out in Archery, with India performing well in Badminton and Athletics. These results highlight the key strengths of each country in specific Olympic sports.

Problem 3 modeling and solution

Model Review

A "great coach"[9] adds h participants per sport, with the model simulating medal effects for three countries using 2024 data. Sensitivity analysis pinpoints coaching-responsive events with high medal gains. Quantitative analysis reveals program contributions, flagging coaching-sensitive ones. Correlation analysis ties medals to factors, using coefficients to spot top-impact programs.

Analysis and Insights

(1) Medal Trends: BP neural network ties 2028 medals to participants, gender, home advantage, and sports. Emerging nations and the U.S. (via home edge) show potential.

(2) "Great Coach" Impact: Sensitivity analysis links coaches to medal gains in judo, gymnastics, and track, e.g., Japan (judo, gymnastics), Mexico (diving), India (track).

(3) Program Correlation: Medal dominance in key sports suggests IOC resource support for fairness and diversity.

Recommendations for the IOC

(1) Resource Allocation: Model guides IOC to fund emerging nations and first-time winners with facilities and coaching.

(2) Program Adjustment: Adjust events per trends, boosting balanced sports for wider investment.
(3) Coaching Investment: "Great coach" analysis directs support, e.g., Japan (judo, gymnastics), Mexico (diving), India (track, shooting).
(4) Data-Driven Strategy: Historical data and updates refine predictions for future Olympics.
Sensitivity Analysis
This study uses a neural network to predict future Olympic medal probabilities, with sensitivity analysis assessing stability, focusing on gender ratio's impact. Key steps:
Step 1: Prep annex data, remove missing values/outliers, train BP Neural Network with historical data, tune for generalization.
Step 2: Select gender ratio and event numbers for analysis.
Step 3: Double these factors, observe prediction shifts.
Sensitivity index quantifies impact, testing U.S. (doubled participants) and Germany (doubled gender ratio), with results in Figure 11:

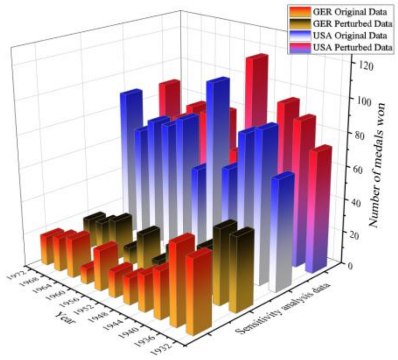

Figure 11. Changes in USA and GER Results for Different Types of Data Perturbations

The model shows high stability and sensitivity through sensitivity analysis. The perturbations to the gender ratio and the number of participants lead to significant changes in the predictions, indicating that the model is sensitive to the response of these key parameters. This indicates that the model has good stability and accuracy in predicting future Olympic medal distribution.
Evaluation of the models
Advantages of the models
(1) High-Precision Prediction: Model forecasts medals using participants, events, and home advantage, predicting Olympic trends accurately.
(2) Multi-Factor Approach: Includes diverse factors like "great coach" effect beyond traditional metrics, assessing impacts holistically.
(3) Key Factor ID: Deep data analysis highlights influences like coaching and specific events, guiding Olympic committees.
(4) Adaptability: Predicts for sports powers and non-winners, offering tailored breakthrough insights.
Disadvantages of the models
(1) Over-simplified Assumptions: Ignoring referee bias or weather may miss key variables, cutting accuracy as chance affects medals.
(2) Data Dependency: Accuracy hinges on data quality; preprocessing can't fully fix missing/outlier issues, risking reliability.
(3) No Dynamic Updates: Lacking real-time adjustments for surprises (e.g., injuries, politics), model flexibility and prediction weaken.
Improvements to the models
(1) Multi-model Integration: To minimize single-model bias, ensemble learning can integrate multiple prediction models, enhancing robustness and accuracy by combining strengths of algorithms like decision trees and support vector machines. (2) Dynamic Adjustment Mechanism: To address future uncertainties, the model can include a dynamic update system, adjusting to new data via continuous real-time input and feedback.
Extension of the models
1) Prediction & Resource Optimization: Model uses historical data to predict Olympic medals, pick events, and allocate resources, optimizing athlete training efficiently.
(2) Coaching & Development: Assesses "great coaching" effect, refining investment and development by tying coaching to performance, targeting key programs, and aiding underperformers.
(3) Cross-Industry Use: Applies to finance (market trends) and medicine (disease forecasts), providing data-driven solutions for efficiency across fields.

3. Conclusion

This study analyzes Olympic medal distributions via predictive modeling and sensitivity analysis to optimize outcomes across countries, providing insights and recommendations for the International Olympic Committee (IOC) and national sports bodies. (1) In Problem One, a BP neural network predicts 2028 medal distributions using factors like athlete participation, gender ratios, home-field advantages, and event correlations, accurately forecasting trends, spotting breakthroughs for less successful nations, and identifying resource allocation opportunities in emerging sports countries. (2) Problem Two explores the "Great Coach" effect, showing how quality coaching boosts medals, with sensitivity analysis highlighting key sports for investment, e.g., judo and gymnastics for Japan, diving for Mexico, and athletics for India, emphasizing coaching's role in medal gains. (3) Problem Three integrates these models' insights, suggesting IOC strategies like resource support for underrepresented nations, prioritizing high-impact sports for coaching, and using data to refine event planning and partnerships. The study links data-driven modeling to strategic sports decisions, offering a framework for competitiveness and inclusivity in the Olympics. Future dynamic modeling with real-time data could enhance adaptability and robustness.

References

[1] Otamendi J ,Doncel M L . Medal Shares in Winter Olympic Games by Sport: Socioeconomic Analysis After Vancouver 2010 [J]. Social Science Quarterly, 2014, 95 (2): 598-614.

[2] Alan V, David H. Prevalence of therapeutic use exemptions at the Olympic Games and association with medals: an analysis of data from 2010 to 2018.[J]. British journal of sports medicine, 2020, 54(15): 920-924.

[3] L. N S .Twin Types: Variations and Common Themes/Twin Research Reviews: Birth Weight and Brain Development; Twinning and Vocabulary Knowledge; Fetal Loss in Twin Pregnancy; Twin-Family Olympic Medal Winners/Human Interest: Young Twin Soldiers; Twin-Run Laundry; Male-Female Pair and Politics; Twin-Based Graphic Novel; Twin Sisters' Deliveries[J]. Twin Research and Human Genetics, 2021, 24(1): 89-94.

[4] Yassin K ,Danilo M, Ervin S .A review of Hidden Markov models and Recurrent Neural Networks for event detection and localization in biomedical signals[J]. Information Fusion, 2021, 6952-72.

[5] Zhiwei G , Lianggui T , Tan G , et al. Deep Graph neural network-based spammer detection under the perspective of heterogeneous cyberspace [J]. Future Generation Computer Systems, 2021, 117205-218.

[6] Zhuhuang Z , Anna G , Weiwei W , et al.Parameter estimation of the homodyned K distribution based on an artificial neural network for ultrasound tissue characterization [J]. Ultrasonics, 2021.

[7] Wang X ,Yao Y ,Li Z , et al. Protocol Reverse Analysis of Ethernet for Control Automation Technology Based on Sequence Alignment and Pearson Correlation Coefficient [J]. Sensors, 2024, 24(24): 7922-7922.

[8] Helmy R ,Steven B , Augie W , et al. Image encoding selection based on Pearson correlation coefficient for time series anomaly detection [J]. Alexandria Engineering Journal, 2023, 82304-322.

[9] Zoltan D. Use one system for all results to avoid contradiction: Advice for using significance tests, equivalence tests, and Bayes factors.[J]. Journal of Experimental Psychology: Human Perception and Performance, 2024,50(5):531-534.