

Together, Individuals Make a Difference**Summary**

Here is the abstract of your paper.

Firstly, that is ...

Secondly, that is ...

Finally, that is ...

$$F(\omega) = \int_{-\infty}^{\infty} f(t) e^{-i\omega t} dt$$

$$f(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} F(\omega) e^{i\omega t} d\omega$$

PCA

%

Keywords: A, B, C,

Contents

1	Introduction	3
1.1	Problem Background	3
1.2	Restatement of Problem	3
1.3	Our work	3
2	Question Preparation	4
2.1	Assumptions	4
2.2	Notations	4
3	Data Preprocessing	4
3.1	Basic Data Preprocessing	4
3.2	Data Mining	6
3.2.1	Athlete Service Status	6
3.2.2	Distribution of Countries' Strength Sports	6
4	Task1:Medal Prediction Model Based on LightGBM	6
4.1	Medal Standings	6
4.2	Countries that Win Their First Medal	9
4.3	Events and Medal Counts by Countries	9
5	Task2:	9
6	Task3	11
7	Task4	12
8	Sensitivity Analysis	12
9	Model Evaluation	12
9.1	Strengths	12
9.2	Weaknesses	12
10	Conclusion	12
	References	12
	Appendix A: Countries' Medal Count Prediction	14
	Appendix B: Program Codes	16

1 Introduction

1.1 Problem Background

1.2 Restatement of Problem

A literature[1] say something about this problem ... There is a conception named "momentum" in tennis, which has a great impact on players' performance. It is a generalization of the influence in manifold aspects like mental stress and residual energy. So the fluctuation of momentum is the most probable factor that reveals the trend of match. However, the momentum is not easy to be quantified for it includes many subjective indicators, and there are few models that can be directly used, so we decide to cut in the following questions from statistical analysis and data procession:

- Develop a model that captures the flow of play as points occur and apply it to one or more of the matches. Your model should identify which player is performing better at a given time in the match, as well as how much better they are performing. Provide a visualization based on your model to depict the match flow. *Note: in tennis, the player serving has a much higher probability of winning the point/game. You may wish to factor this into your model in some way.*
- A tennis coach is skeptical that "momentum" plays any role in the match. Instead, he postulates that swings in play and runs of success by one player are random. Use your model/metric to assess this claim.
- Coaches would love to know if there are indicators that can help determine when the flow of play is about to change from favoring one player to the other.
 - 1) Using the data provided for at least one match, develop a model that predicts these swings in the match. What factors seem most related (if any)?
 - 2) Given the differential in past match "momentum" swings how do you advise a player going into a new match against a different player?
- Test the model you developed on one or more of the other matches. How well do you predict the swings in the match? If the model performs poorly at times, can you identify any factors that might need to be included in future models? How generalizable is your model to other matches (such as Women's matches), tournaments, court surfaces, and other sports such as table tennis.
- Produce a report of no more than 25 pages with your findings and include a one- to two-page memo summarizing your results with advice for coaches on the role of "momentum", and how to prepare players to respond to events that impact the flow of play during a tennis match.

1.3 Our work

We do such things ...

1. We do ...
2. We do ...
3. We do ...

2 Question Preparation

2.1 Assumptions

2.2 Notations

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

Table 1: Notations

Symbol	Definition
C_i	the first one
S_i	the second one
G	the second one
V	the second one
B	the second one
M	the second one
A	the second one
D	the second one
E	the second one
γ	Sport Advantagen Coefficient
m_{c_i}	the second one
$m_{c_i s_j}$	the second one
g_i	the second one
ξ	The Probability of First Medal

3 Data Preprocessing

3.1 Basic Data Preprocessing

$$G_i \cap G_j = \emptyset, \bigcup_{i=1}^5 G_i = X$$

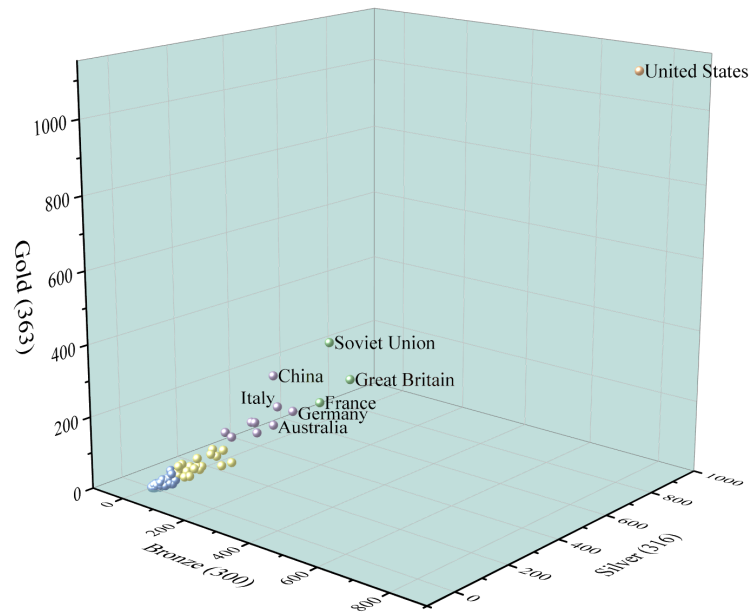


Figure 1: Scatter plot of national level classification (based on Kmeans++clustering algorithm)

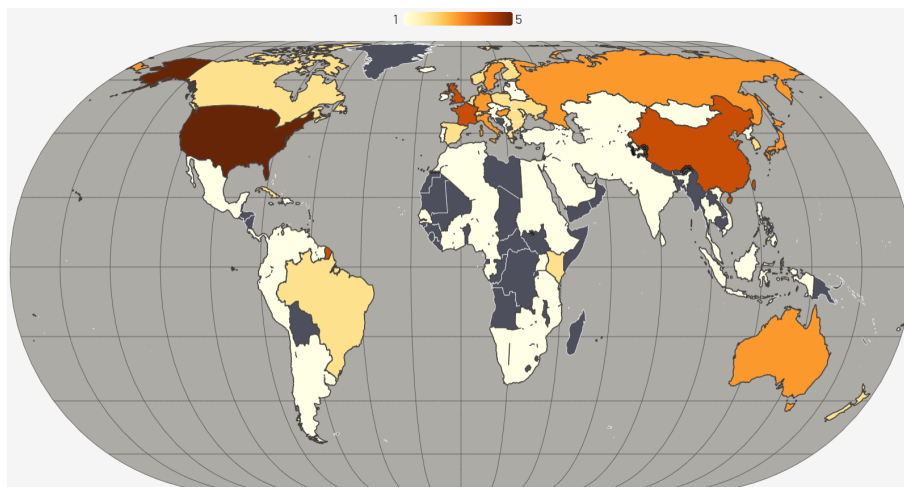


Figure 2: aa

3.2 Data Mining

3.2.1 Athlete Service Status

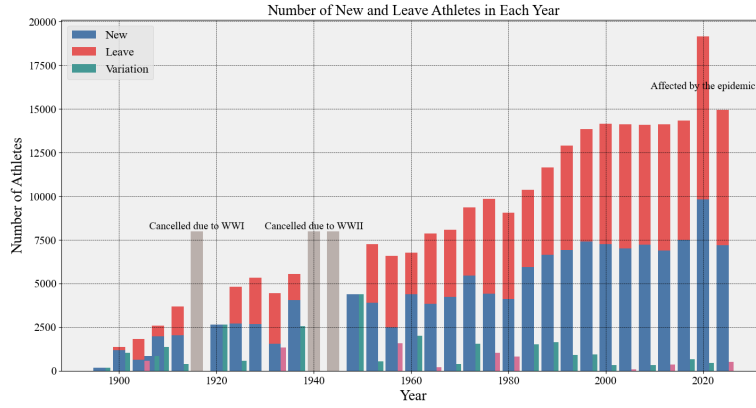
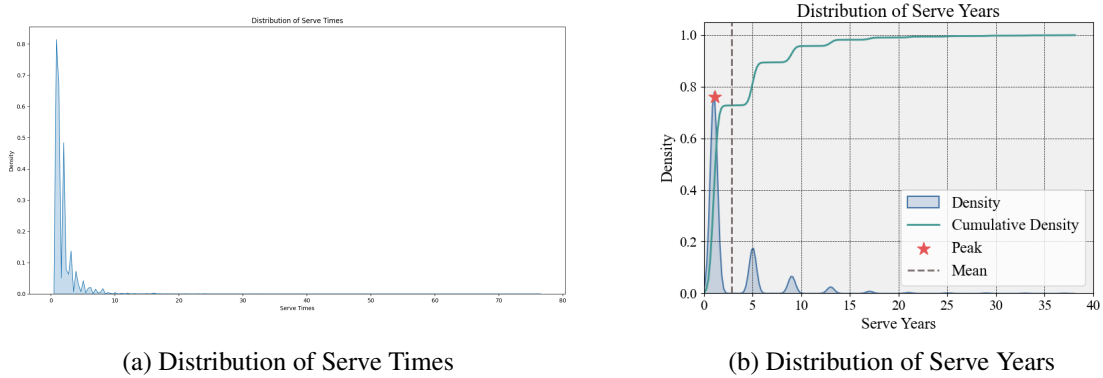


Figure 3: Number of New and Leave Athletes in Each Year



(a) Distribution of Serve Times

(b) Distribution of Serve Years

Figure 4: Two images

3.2.2 Distribution of Countries' Strength Sports

4 Task1:Medal Prediction Model Based on LightGBM

4.1 Medal Standings

The detail can be described by equation (1):Equation 1:

$$\gamma = \frac{m_{c_i s_j}}{\sum m_{c_i s_j}} \quad (1)$$

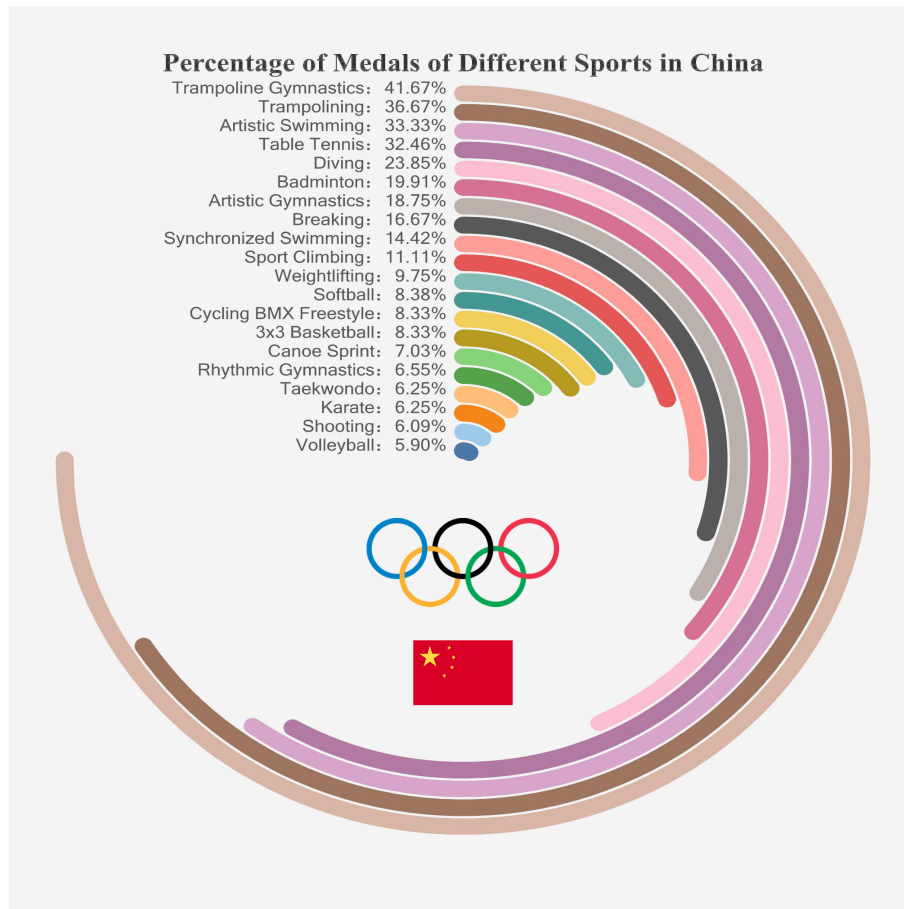


Figure 5: Percentage of Medals of Different Sports in China

Table 2: Variable Name

Variable Name	Code	Definition
Whether Host Country	<i>is host</i>	Whether the country is the host(1 for host,0 for non-host)
Medal Expectation Increment *Personnel Expectation Increment	<i>medal_increment * personnel_increment</i>	Product of medal expectation increment and personnel expectation increment
Sport Advantage Coefficient	<i>sport_adv</i>	Advantage coefficient of a specific sport
Country Level	<i>country_lvl</i>	The level of the country in the competition (ordered by rank)
Project Medal Expectation /Project Personnel Expection	<i>sport_medal _per_ person</i>	Ratio of sport medals to projected personnel for a specific sport
Gold Medal Probability	<i>gold_prob</i>	Probability of an athlete winning a gold medal

Continued on next page

Table 2: Variable Name (Continued)

Silver Medal Probability	<i>silver_prob</i>	Probability of an athlete winning a silver medal
Bronze Medal Probability	<i>silver_prob</i>	Probability of an athlete winning a bronze medal
No Medal Probability	<i>no_medal_probe</i>	Probability of an athlete winning no medal

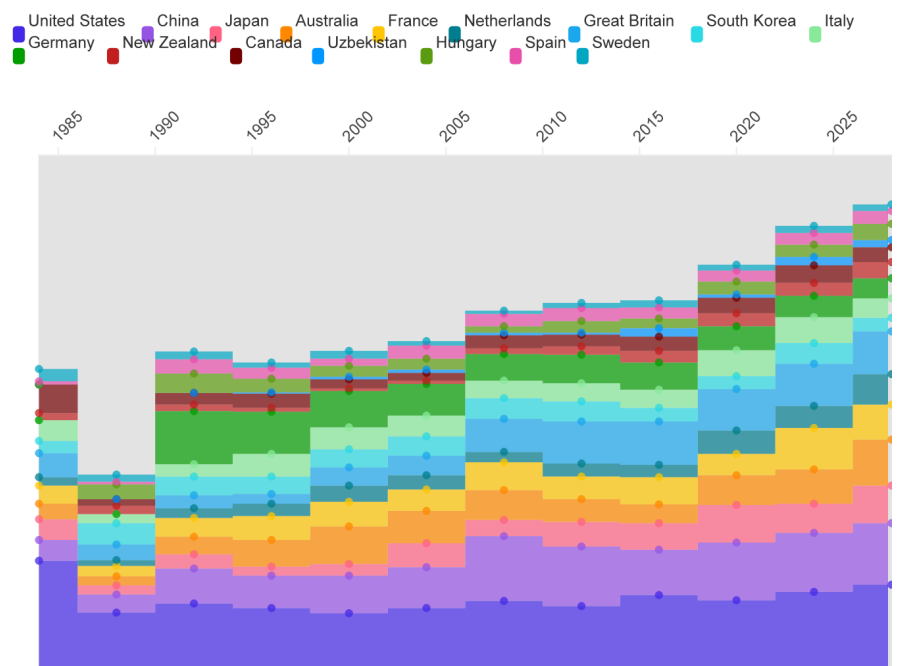


Figure 6: Medal count prediction

Table 3: Countries' Medal Count Prediction(part)

Country	G_ pes.	S_ pes.	B_ pes.	G_ opt.	S_ opt.	B_ opt.	Gold	Silver	Bronze	Total
United States	47	36	23	53	41	25	51	40	25	117
China	35	35	15	46	38	20	40	36	17	95
Australia	28	20	22	33	24	24	29	22	23	71
France	23	19	19	27	20	22	24	19	21	64
Germany	18	24	14	23	26	16	21	25	15	61
United Kingdom	20	19	14	26	24	15	25	21	15	61
Japan	16	25	14	19	27	16	17	26	16	58
Italy	15	20	15	17	20	22	16	20	16	53
Spain	21	10	13	30	12	14	25	12	13	50
Netherlands	24	6	13	29	9	15	26	8	15	47

New Zealand	20	13	9	22	15	10	22	14	9	45
...

4.2 Countries that Win Their First Medal

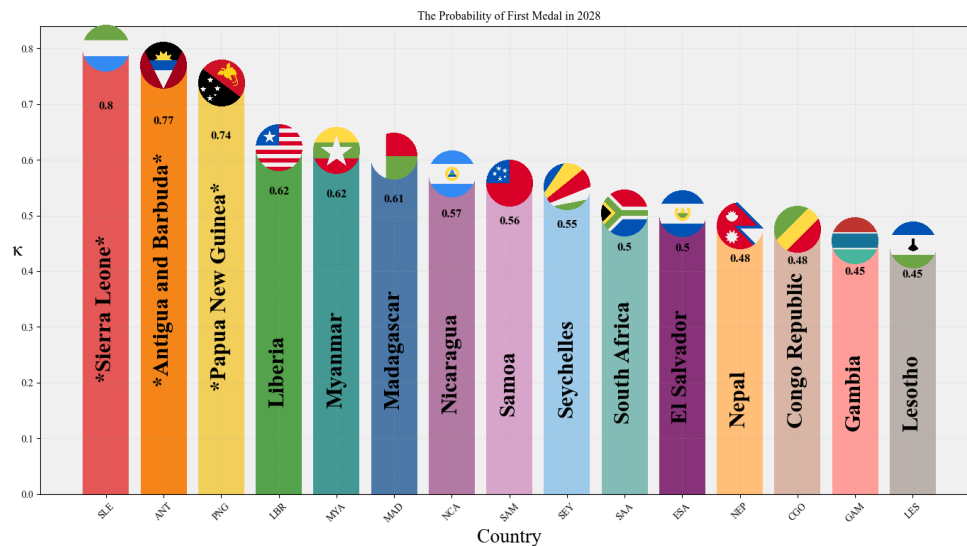
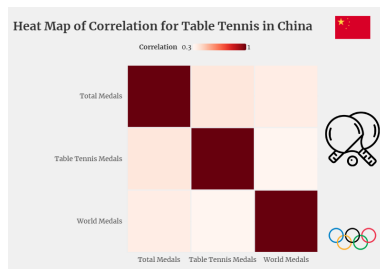


Figure 7: The Probability of First Medal in 2028

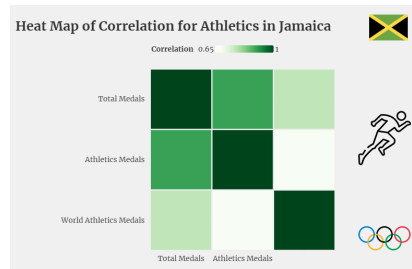
4.3 Events and Medal Counts by Countries

5 Task2:

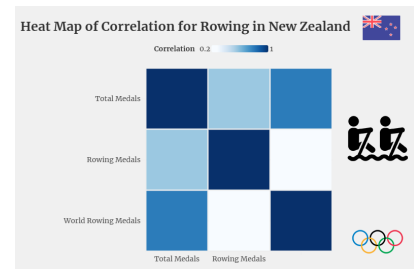
The results are shown in Figure ??, where t denotes the time in seconds, and c refers to the concentration of water in the boiler.



(a) Heatmap of Correlation for Table Tennis in China



(b) Heatmap of Correlation for Athletics in Jamaica



(c) Heatmap of Correlation for Rowing in New Zealand

Figure 8: Three images

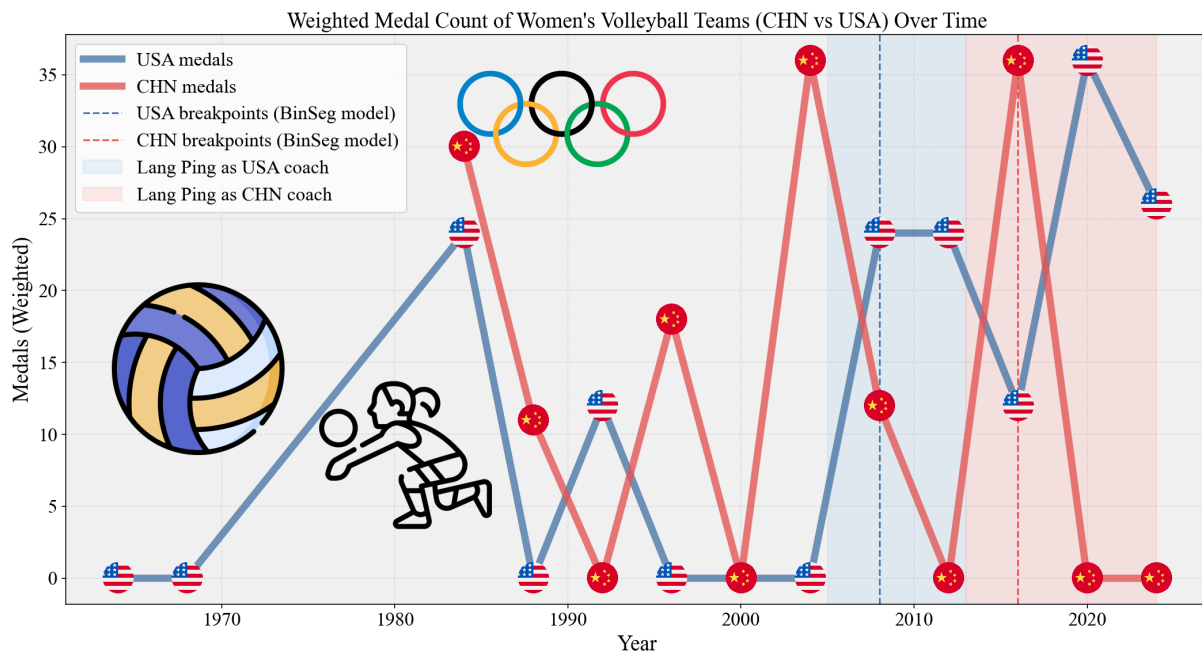
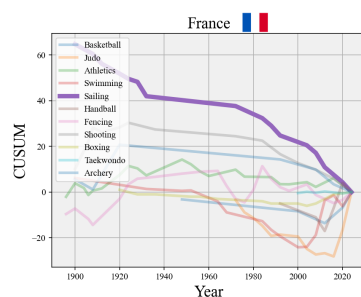
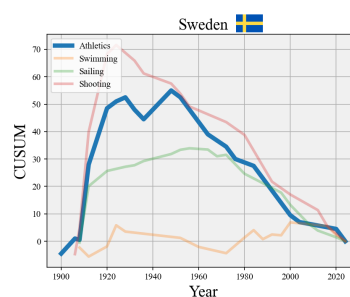


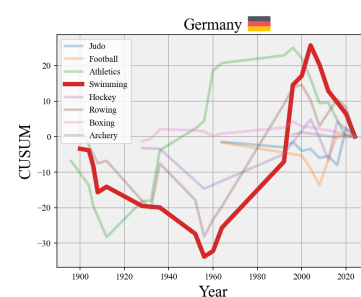
Figure 9: Medal Count of Women's Volleyball Teams(CHN vs USA) Over Time



(a) France



(b) Sweden



(c) Germany

Figure 10: Three images

Table 4: Model Performances(LightGBM)

Model	MSE	RMSE	MAE	R ²
Gold Prediction	0.0235	0.180	0.030	0.690
Silver Prediction	0.0239	0.185	0.033	0.625
Bronze Prediction	0.0231	0.187	0.034	0.606

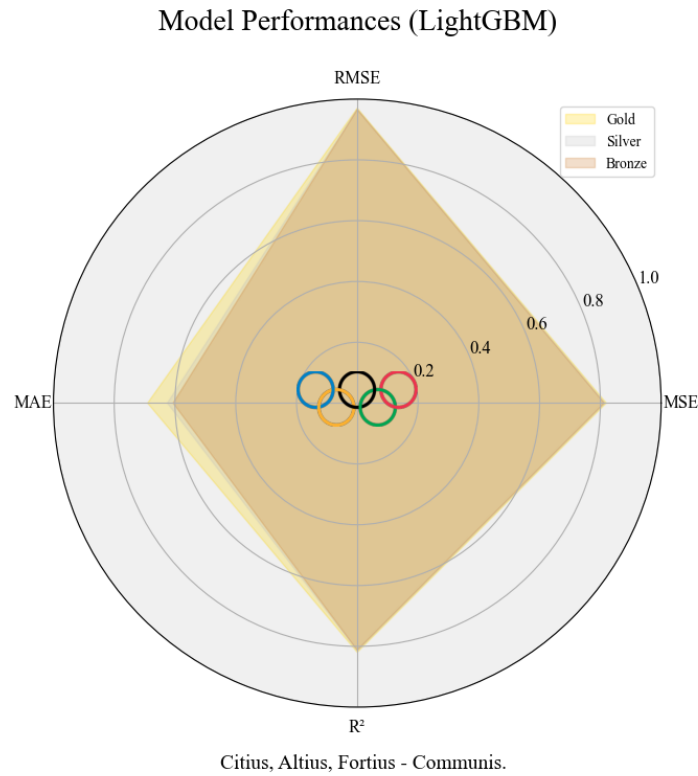


Figure 11: Model Performance Radar Chart

6 Task3

$$\mathcal{L} = \sum_{i=1}^n L(y_i, \hat{y}_i^{(m-1)} + \eta f_m(x_i)) + \Omega(f_m) \quad (2)$$

$$\begin{aligned} & y_i \\ & \hat{y}_i^{(m-1)} \\ & \eta f_m(x_i) \\ & \Omega(f_m) \end{aligned}$$

$$\kappa = 0.8 \times \frac{\xi - \bar{\xi}}{\sigma} \quad (3)$$

$$\bar{\xi} \sigma$$

$$d_i = R(X_i) - R(Y_i) \quad (4)$$

$$R_s = 1 - \frac{6 \sum d_i^2}{n(n^2 - 1)} \quad (5)$$

$d_i R_s$

$$P(c_k | X) = \frac{P(X | c_k)P(c_k)}{P(X)} \quad (6)$$

$$C_t^+ = \max(0, C_{t-1}^+ + (X_t - \mu - k)) \quad (7)$$

C_t^+

7 Task4

8 Sensitivity Analysis

9 Model Evaluation

9.1 Strengths

- 1.
- 2.
- 3.
- 4.

9.2 Weaknesses

- 1.

10 Conclusion

- Task1
- Task2

- Task3
- Task4

References

- [1] Einstein, A., Podolsky, B., & Rosen, N. (1935). Can quantum-mechanical description of physical reality be considered complete?. *Physical review*, 47(10), 777.
- [2] *A simple, easy L^AT_EX template for MCM/ICM: EasyMCM*. (2018). Retrieved December 1, 2019, from <https://www.cnblogs.com/xjtu-blacksmith/p/easymcm.html>

Appendix A: Countries' Medal Count Prediction

Table 5: Countries' Medal Count Prediction

Country	G_ pes.	S_ pes.	B_ pes.	G_ opt.	S_ opt.	B_ opt.	Gold	Silver	Bronze	Total
United States	47	36	23	53	41	25	51	40	25	117
China	35	35	15	46	38	20	40	36	17	95
Australia	28	20	22	33	24	24	29	22	23	71
France	23	19	19	27	20	22	24	19	21	64
Germany	18	24	14	23	26	16	21	25	15	61
United Kingdom	20	19	14	26	24	15	25	21	15	61
Japan	16	25	14	19	27	16	17	26	16	58
Italy	15	20	15	17	20	22	16	20	16	53
Spain	21	10	13	30	12	14	25	12	13	50
Netherlands	24	6	13	29	9	15	26	8	15	47
New Zealand	20	13	9	22	15	10	22	14	9	45
Canada	15	10	16	19	10	18	17	10	16	43
Brazil	8	10	13	8	11	15	8	10	14	32
Belgium	6	8	7	15	10	8	11	8	7	26
Hungary	7	6	8	8	8	9	8	8	9	25
Poland	10	7	6	12	10	8	11	9	6	25
Ireland	6	10	6	7	11	6	6	10	6	23
Argentina	4	8	7	5	10	9	4	9	8	22
Denmark	10	7	5	12	8	8	10	8	6	22
Ukraine	3	7	8	5	8	9	4	8	8	21
South Korea	10	5	4	12	6	4	11	6	4	21
Romania	10	7	3	12	8	3	12	7	3	20
Norway	11	1	4	11	3	6	11	2	6	19
South Africa	6	4	4	8	8	5	8	7	5	18
Slovenia	4	3	8	6	3	8	4	3	8	17
Kenya	5	4	4	5	7	5	5	7	4	16
Serbia	4	5	6	5	9	7	4	8	6	16
Mexico	4	4	3	6	4	5	5	4	4	15
India	1	1	4	8	4	5	7	3	4	15
Greece	3	4	2	7	4	3	7	4	3	14
Switzerland	2	4	5	5	4	6	4	4	6	14
Jamaica	1	7	2	5	12	3	4	8	2	13
Czech Republic	5	3	3	6	4	5	6	4	4	13
Uzbekistan	4	2	2	5	3	2	4	2	2	10
Nigeria	4	2	4	4	3	4	4	2	4	10
Sweden	4	4	2	5	5	5	4	4	2	10
Turkey	3	2	5	3	2	6	3	2	5	10
Finland	2	6	2	2	6	2	2	6	2	10
Colombia	5	2	0	5	3	1	5	3	0	9

Egypt	0	3	3	1	4	4	0	4	4	8
Kazakhstan	1	2	2	4	4	2	3	3	2	8
Iran	2	2	1	5	3	2	4	2	1	8
Dominican Republic	3	3	2	6	4	3	5	3	3	8
Portugal	2	3	3	2	4	3	2	3	3	8
Cuba	2	2	1	3	2	2	3	2	2	7
Unknown	1	2	2	1	2	4	1	2	2	7
Latvia	3	1	3	3	1	3	3	1	3	7
Ethiopia	2	1	0	2	6	2	2	2	2	6
Algeria	2	2	1	2	3	2	2	2	1	6
Morocco	0	2	4	0	2	5	0	2	5	6
Austria	3	0	1	3	1	2	3	0	2	6
Puerto Rico	2	3	1	4	3	2	2	3	2	6
Thailand	1	2	1	2	2	2	2	2	1	5
Croatia	2	2	0	2	2	1	2	2	1	5
Ecuador	2	0	2	3	1	3	2	0	3	5
Uganda	0	1	1	1	4	1	1	3	1	5
Angola	0	2	0	1	3	1	0	2	0	5
Israel	2	2	1	2	2	2	2	2	1	5
Mongolia	1	3	1	1	4	1	1	4	1	5
Fiji	0	1	0	2	5	1	1	4	0	5
Uruguay	0	2	0	1	3	1	0	2	0	5
Chile	0	2	0	1	3	1	0	2	0	5
Unknown	1	1	0	3	1	0	1	1	0	4
Zambia	1	3	0	1	3	0	1	3	0	4
Azerbaijan	1	0	2	1	1	2	1	1	2	4
Guinea	0	4	0	0	4	0	0	4	0	4
Georgia	1	1	0	2	2	0	2	2	0	3
Mali	0	3	0	0	3	0	0	3	0	3
Chinese Taipei	2	1	0	2	1	0	2	1	0	3
Venezuela	0	2	1	1	2	2	0	2	2	3
Paraguay	0	3	0	0	4	0	0	4	0	3
Indonesia	0	1	1	1	1	2	1	1	2	3
Peru	0	1	0	1	2	0	0	2	0	3
Iraq	0	3	0	0	3	0	0	3	0	3
Bulgaria	1	1	0	1	1	1	1	1	0	3
Lithuania	2	1	0	2	2	1	2	1	1	3
Tunisia	0	0	0	1	1	0	0	0	0	2
Guatemala	1	0	1	1	0	1	1	0	1	2
Bahamas	0	2	0	1	2	0	0	2	0	2
Trinidad and Tobago	0	1	0	1	1	0	0	1	0	2

Montenegro	0	1	0	0	2	0	0	2	0	2
Hong Kong	0	1	1	0	1	1	0	1	1	2
Moldova	0	1	0	0	2	0	0	2	0	1
South Sudan	0	1	0	0	2	0	0	2	0	1
Malaysia	0	1	0	0	1	0	0	1	0	1
Philippines	0	0	0	1	0	0	0	0	0	1
Bahrain	0	0	0	0	2	0	0	1	0	1
Estonia	0	0	1	0	0	1	0	0	1	1
Liberia	0	1	0	0	2	0	0	1	0	1
United Arab Emirates	0	0	0	0	1	0	0	0	0	1
Jordan	0	0	0	0	1	0	0	0	0	1
Eritrea	0	1	0	0	2	0	0	2	0	1
Singapore	0	0	0	0	1	1	0	0	1	1
Luxembourg	0	0	0	0	1	0	0	0	0	1
Cyprus	0	0	0	0	1	0	0	0	0	1
Samoa	0	1	0	0	1	0	0	1	0	1
Qatar	0	0	0	0	0	0	0	0	0	0
Tanzania	0	0	0	0	0	0	0	0	0	0
Djibouti	0	0	0	0	0	0	0	0	0	0
Ghana	0	0	0	0	0	0	0	0	0	0
Ivory Coast	0	0	0	0	0	0	0	0	0	0
Tajikistan	0	0	0	0	0	0	0	0	0	0
Kyrgyzstan	0	0	0	0	1	0	0	0	0	0
Armenia	0	0	0	0	0	0	0	0	0	0
Slovakia	0	0	0	0	0	0	0	0	0	0
Botswana	0	0	0	0	2	0	0	0	0	0
Grenada	0	0	0	0	0	0	0	0	0	0
Zimbabwe	0	0	0	0	0	0	0	0	0	0
Burundi	0	0	0	0	0	0	0	0	0	0
Kosovo	0	0	0	0	0	0	0	0	0	0

Appendix B: Program Codes

Here are the program codes we used in our research.

test.py

```
# Python code example
for i in range(10):
    print('Hello, world!')
```

test.m

```
% MATLAB code example
```



```
for i = 1:10
    disp("hello, world!");
end
```

test.cpp

```
// C++ code example
#include <iostream>
using namespace std;

int main() {
    for (int i = 0; i < 10; i++)
        cout << "hello, world" << endl;
    return 0;
}
```