| Problem Chosen | 2021 MCM/ICM Summary Sheet | Team Control Number |
|---|---|---|
| F | | 2120850 |

# Healthy and Sustainable Higher Education System


Nowadays with the development of economic and technological globalization, the higher education environment is gradually developing rapidly. Although under the influence of different domestic political environments, each country has its own unique measuring standards for higher education, but almost every country paid more and more attention to the higher education system, and realized the importance of high-quality higher education from nature, especially for economic prosperity, scientific and technological progress and the improvement of citizen welfare. Therefore, discovering the way to make this system healthier and more sustainable is now demonstrating its importance and urgency, and it will be quite challenging in future implementation.

This paper will establish a scientific evaluation and prediction model in this article to evaluate the health and sustainability of higher education system across different nations. It includes four basic models: PCA Model, Grey Prediction Model, ARIMA Prediction Model and BP&PSO Advanced Neural Network Model. They solved the 7 sub-problems of the problem.

In the process of model building and problem solving, there are mainly four steps. First, collect and process effective data, then perform Principal Component Analysis on the original data to reduce its dimension. Second, delimit the three principal components after data feature extraction, then use the grade evaluation method to grade them in sequence. Next, establish The BP Neural Network optimized by Particle Swarm Optimization to train the processed data. We can obtain an effective higher education health evaluation model after training. Third, the Grey-ARIMA hybrid prediction method can reduced the error of the forecast data. It forward processes the difference between the forecast data and the current data, then the BP neural network is trained again with these data, finally the sustainable model of higher education is established precisely. Fourth, a set of established models were applied to a wider range of countries to evaluate their rationality. After the overall evaluation, This paper selects and analyzes the areas that need to be improved in Brazil's higher education, and sets up a timetable to propose policy recommendations and discuss their feasibility.

This model has passed the sensitivity test, and has certain generalization in various fields. At the same time, we also proposed to improve the neural network model by learning rate attenuation, modifying the node number of hidden layer, using momentum gradient descent method and other ways to effectively improve the convergence speed and solution accuracy.



**Keywords**: Higher Education, PCA Model, Grey Prediction Model, ARIMA Prediction Model, BP Neural Network Model, Particle Swarm Optimization

# Contents

# 1. Restatement of Problem

## 1.1 Background

Higher education is a non-compulsory stage after high school, including vocational education and professional education. The level of higher education is an important indicator to measure the comprehensive strength of a country. However, different countries have different priorities[1] for higher education. For example, the United States focuses more on liberal arts education, while the United Kingdom focuses more on science and engineering education. In recent years, due to the continuous improvement of international science and technology, the demand for high-level talents in various countries has also increased. In the field of higher education research, how to make the country's education develop in a healthier direction has always been a hot topic. Therefore, it is the challenge of this paper to establish a fair, healthy and sustainable higher education system which can be adapted to the evaluation of the education level of all countries around the world.

## 1.2 Problems Analysis

This topic requires the development and verification of a model or set of models that can be used to evaluate the health of the higher education system in any country. Based on the analysis, we select a country with room for improvement, put forward more reasonable suggestions for improvement based on its current healthy and sustainable system status, and predict whether it can achieve the expected vision through the model, finally, the realistic impact of the implementation plan will be discussed according to the policy.

Based on the above problem analysis, the author will solve this problem through four steps.

1.Develop a model to evaluate the health and sustainability of the country's higher education system, and validate the model with multi-country data.

2. After applying the model to several other countries, based on the results of data analysis, select a country with room for improvement in the higher education system, analyze the health and sustainability of the selected country, then, propose an achievable and reasonable vision.

3. Propose targeted policies and implementation timetables for the country, based on specific indicator analysis, to support development from the current state to the proposed state, using the model to develop and/or assess the effectiveness of the policies.

4. Discuss the practical impact of implementing the plan during the transition and final stages (for example, the impact on students, teachers, schools, communities, and the country).

### 1.3 Interpretation of Terms

**Higher Education**: an optional stage of formal learning that occurs after completion of secondary level of education, including graduate, master, doctor, etc.

**System of Higher Education**: an organizational structure that consists of higher educational institutions as well as personnel and infrastructure required to educate students beyond the secondary level.

**Sustainable System**: a system that maintains its effectiveness over time.

**System Health**: a system to align around a common vision, execute against that vision effectively, and renew itself through innovation and creative thinking.

# 2. Model Hypothesis and Symbol Description

## 2.1 Model Assumptions

1.Assuming that the data about countries collected from major websites are true and reliable.

2. Assuming that the 12 indicators in the article can describe the level of higher education.

3.Assuming that it is reliable to use the higher education level of five countries as the training set of the neural network.

## 2.2 Symbol Description

Table 2.2.1 Symbols Table

| Serial number | Symbol | Symbols indicate meaning |
|---|---|---|
| 1 | $X$ | n-dimensional random vector |
| 2 | $M_m$ | M-th principal component |
| 3 | $X_n$ | Nth index |
| 4 | $p$ | Autoregressive order |
| 5 | $d$ | Difference order |
| 6 | $q$ | Moving average order |
| 7 | $I_j$ | Node input |
| 8 | $O_j$ | Actual output value |

| | | |
|---|---|---|
| 9 | $W_{ij}$ | Weight between the two layers before and after |
| 10 | $\theta_j$ | Threshold between two layers |
| 11 | $lr$ | Learning rate |

# 3. Establishment of Basic Models

## 3.1 Data Collection and Processing

This paper collected more than 30 indicators describing the higher education system through World Bank, Human Development Report, OCED, The National Science Foundation's National Center for Science and Engineering Statistics, and included 300 countries around the world. In order to reflect the level of global higher education at various levels, this paper selected data from seven countries: Australia, Brazil, France, Turkey, the United States, Germany, and Japan. Due to the incompleteness of some data and the subsequent prediction model, this paper selected data from seven countries. For data processing and screening, this paper finally selected 12 indicators that affect the level of higher education from 2014 to 2018. These 12 indicators can summarize a healthy and sustainable higher education system in all aspects from the perspectives of input cost, degree value, scientific research level, and education quality. The following is a specific description of each indicator.

The average years of education of 25 years and above (It is referred to as AYE) can not only measure the level of the national education system, but also represent the level of modernization; the average mathematics and science scores (AMS) of countries before higher education can reflect the pre-higher education Scientific level; Science & Engineering articles (SEA) in all fields of each country can indicate the level of science and technology in all fields of the country, including science, engineering, agriculture, medicine and other important subjects; the number of higher education students studying abroad (SSA) reflects part The proportion of higher education going to other countries; the global higher education enrollment rate (percentage) (GER) and the total graduation rate of higher education bachelor degree (TGR) in various countries can be analyzed together to show how difficult it is to complete higher education in each country, The number of international students (INS) can reflect from the side how many international students a country's educational level can attract; government expenditure on education (GEF), per capita government expenditure on higher education students(PEE), and national R&D expenditure (percent of GDP)

(NRE) It can reflect the importance of the country for higher education and high-level scientific research; the gender parity index (IGE) for higher education enrollment reflects the degree of gender equity in higher education to a certain extent; the youth literacy rate (YLR) can be reflected in the humanities The degree of popularization of national education.

## 3.2 Basic Models

In order to establish a set of models for measuring the health and sustainability of higher education, this article uses the following four basic models, namely Principal Component Analysis Model, Grey-ARIMA Hybrid Prediction Model, BP Neural Network Model and BP&PSO Advanced Neural Network Model The basic relationships and processes of the four models are shown in Figure 3.2.1.1.

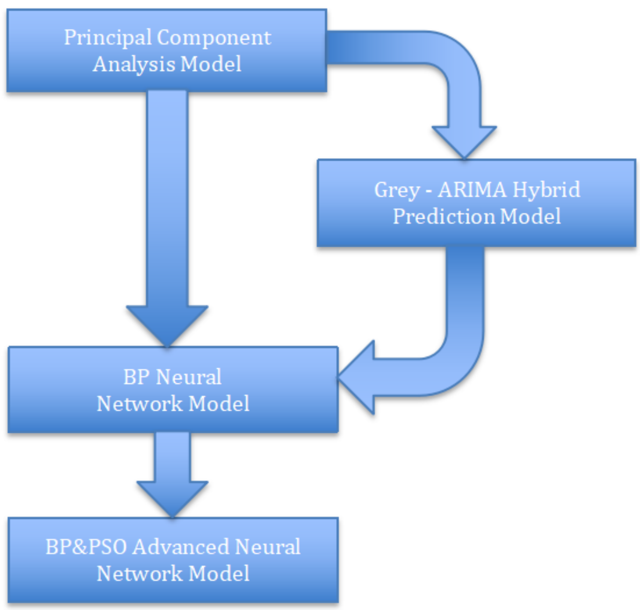


Figure 3.2.1 The Relationships and Processes of Four Models

### 3.2.1 Principal Component Analysis

When establishing the model of higher education system, the author collected a large number of relevant characteristic attributes, then screened out 12 effective indicators. The more metrics collected, the more comprehensive higher education system, but the more data also need to deal with. And there are correlation between many indicators, so need to dimension of index, so we established the basic model of principal component analysis (PCA), The n-dimensional index is mapped to the m-dimensional index, that is, the m-dimensional index is used to replace the original n-dimensional index, and the m-dimensional index is called the main component. Principal component can not only replace the original index to describe the higher education model completely, but also reduce the calculation of data.

$$X = (X_1, X_2, X_3, \cdots, X_n) \quad (1)$$

$$M_1 = a_1' X = a_{11} X_1 + a_{21} X_2 + \cdots + a_{n1} X_n \quad (2)$$

$$M_2 = a_2' X = a_{12} X_1 + a_{22} X_2 + \cdots + a_{n2} X_n \quad (3)$$

$$M_3 = a_3' X = a_{13} X_1 + a_{23} X_2 + \cdots + a_{n3} X_n \quad (4)$$

……

$$M_m = a_m' X = a_{1n} X_1 + a_{2n} X_2 + \cdots + a_{nn} X_n \quad (5)$$

(1)X is an n-dimensional random vector.

(2)-(5)In principal component analysis, m-dimensional vector is used to replace the original n-dimensional vector.

After establishing the basic model, we performed KMO and Bartlett tests on 12 indicators to observe whether they are suitable for principal component analysis. The KMO value calculated by SPSS is 0.675, and the Bartlett sphericity test value is 45 ($p<0.05$), indicating that the index is suitable for the principal component analysis method to be carried out next.

**3.2.2 Back Propagation Neural Network and Particle Swarm Optimization**

**3.2.2.1 Overview**

Back Propagation neural network is a traditional neural network model. By training it with an effective and appropriate data set, the model can infer the most likely target solution based on the new data. This kind of neural network usually includes an input layer, a hidden layer and an output layer. The input layer is the independent variable necessary to obtain the target solution, and the output layer is the corresponding target solution. The correctness of this target solution depends on the selection of the data set. And training batches. In addition, what is contained between the input layer and the hidden layer is the weight. This weight determines the importance of a certain input variable in this solution. After a series of transformations and calculations in the hidden layer, it will finally be hidden The data transmitted by the layer is connected with the final output layer, and the final prediction result is output after classification. The process of Back Propagation neural network establishment is shown in Figure 3.2.2.1.

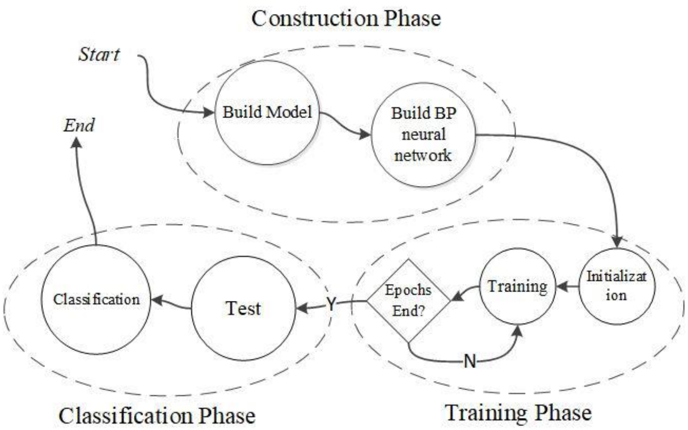


Figure 3.2.2.1.1 Construction Phase

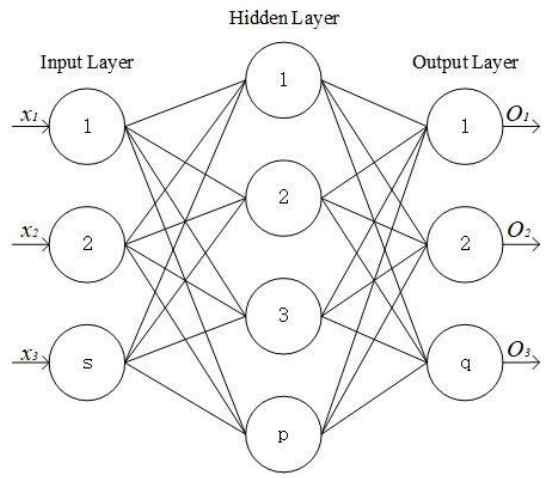


Figure 3.2.2.2.2 Hidden Layer

### 3.2.2.2 Back Propagation Neural Network Principle

1. Initialization

The initialization process of Back Propagation neural network needs to determine the content and number of nodes *s* in the input layer, the number of nodes *p* in the hidden layer, the learning rate $\eta$, the initial value of momentum *n*, and the number of nodes *q* in the output layer.

2. Forward propagation

Starting from the input layer, we must start forward propagation, that is, calculate the input $I_j$ and actual output value $O_j$ of node j in the next layer from front to back. $W_{ij}$ is the weight between the two layers before and after, and $\theta_j$ is the threshold between the two layers, The forward propagation formula is as follows.

$$I_j = \sum_i W_{ij} O_i - \theta_j \tag{6}$$

$$O_j = 1/(1+e^{-I_j}) \tag{7}$$

3. Calculate the error

In the training process, if the target solution does not match the actual value, a backpropagation algorithm will be called to further modify the previous weights. The error E needs to be calculated for the output layer node j, and $T_j$ is the expected output value. Then start the reverse broadcast according to this error value, the error calculation formula is as follows.

$$E_j = O_j(1-O_j)(T_j - O_j) \tag{8}$$

4.Backpropagate and update weights and thresholds

$$E_j = O_j(1 - O_j)\sum_k E_k W_{jk} \tag{9}$$

$$\Delta W_{ij} = \eta E_j O_j \tag{10}$$

$$W_{ij} = W_{ij} + \Delta W_{ij} \tag{11}$$

$$\Delta \theta_j = \eta E_j \tag{12}$$

$$\theta_j = \theta_j + \Delta \theta_j \tag{13}$$

(9) is the back propagation formula, (10)-(13) is the weight and threshold update formula.

#### 3.2.2.3 Particle Swarm Algorithm Optimizes Neural Network

The PSO algorithm, that is, the particle swarm optimization algorithm, can well assist in optimizing neural networks. This algorithm first initializes a group of particles in an N-dimensional solution space. Here, the number of particles is set to n, and the position of a single particle in the solution space D is set to $X_i$, and its velocity is $V_i$, but in order to prevent the particle's speed from being too fast, $V_i$ must not exceed a certain set maximum speed value. Similarly, a single particle cannot run randomly, and its position $X_i$ must be limited to a constraint function. It is easy to know that each particle has the possibility of becoming the optimal solution, and its fitness value can be obtained by substituting $X_i$ into the objective function, and the positions of the best fitness values of individual particles and group particles are set to Bi and Ba, here at the same time, the position of the particles in the particle swarm is constantly updated, and the update formula is as follows:

$$V_i = w \times V_i + c1 \times r1 \times (B_i - X_i) + c2 \times r2 \times (Ba - X_i) \tag{14}$$

$$X_i = X_i + V_i \tag{15}$$

(15) w is the inertia factor, and $w \geqslant 0$ is specified. c1 and c2 are defined as acceleration constants, r1 and r2 are random numbers, which stipulate that $r1, r2 \in [0,1]$.

### 3.2.3 Grey-ARIMA Hybrid Prediction Model[2]

The Grey Prediction Model is to predict the future development of things by establishing corresponding differential equations. It is not only suitable for forecasting large amounts of data, but also effective forecasting for a small number of sequences. This paper used the GM(1,1) model, which is a first-order differential equation model containing a sequence of numbers.

$$\hat{x}^{(1)}(k+1)=(x^{(0)}(1)-\frac{\hat{b}}{\hat{a}})e^{-\hat{a}k}+\frac{\hat{b}}{\hat{a}},k=0,1,\cdots n-1,\cdots \tag{16}$$

a is the development coefficient, b is the gray effect quantity. a and b are the values used in the construction of the model, which have little practical significance. The posterior difference ratio C value 0.231<=0.35, which means that the model accuracy level is very good.

ARIMA model is a model of time series forecast analysis. Among them, we can split it into three parts.

$$X_t=\alpha_1 X_{t-1}+\alpha_2 X_{t-2}+\cdots+\alpha_p X_{t-p}+u_t \tag{17}$$

$$u_t=\varepsilon_t+\beta_1\varepsilon_{t-1}+\cdots+\beta_q\varepsilon_{t-q} \tag{18}$$

(17) is the AR (Auto Regression) model.

(18) is the MR (Moving Average) model, $\varepsilon_t$ is the white noise sequence.

By combining the autoregressive model (AR), the moving average model (MA) and the difference method, we get the differential autoregressive moving average model ARIMA (p, d, q).Where p is the autoregressive order, and d is the order that needs to be differed on the data. q is the moving average order.

Both the grey prediction model and the ARIMA model have some limitations and shortcomings. For example, although the grey prediction can fit the data of the next 12 periods, only the short-term prediction is more reliable. Therefore, the author adopts Grey-ARIMA Hybrid Prediction Model, which can make the advantages and disadvantages of the two prediction models complement each other and achieve the maximum prediction accuracy.

# 4. Solution of the Problem

## 4.1 Establishment of a Healthy and Sustainable Model of the Higher Education System

### 4.1.1 General Description

In this part, a set of model will be developed to evaluate the health and sustainability of the country's higher education system, and verify this model with data from multiple countries. This set of model includes four basic models: PCA Model, Grey Prediction Model, ARIMA Prediction Model and BP&PSO Advanced Neural Network Model. Among them, PCA method reduces the dimensionality of 12 indicators and obtains three principal components. The three principal components are used as the input layer and passed into the BP Neural Network, and the output layer is the corresponding. The goal solution is the health of the higher education system. The Grey-ARIMA hybrid prediction method to predict the values of the three principal components in the future can reduced the error of the forecast data, then the values re-enter the Neural Network to output the sustainability of the higher education system.

**4.1.2 Principal Component Analysis Method for Dimensionality Reduction**

After the test by KMO and Bartlett, the author used SPSS software to standardize the original data, and then carried out principal component analysis on these data.

First, according to the variance of the common factor and the size of the load coefficient, two unreasonable indicators were eliminated. Among them, the common degree (variance of common factor) value of SER is 0.116 and less than 0.4, which means that this index has little correlation with all principal components, so it should be eliminated. The absolute value of the load coefficient of GSR is less than 0.4 under the three principal components, so this index also needs to be deleted.

After the deletion, there are only 10 indicators remaining. After that, the process of data concentration is to summarize the original 10 indicators into 3 principal components. The characteristic root values of these three principal components are all more than one, the variance explanation rates are 60.219%, 19.125%, and 10.827%, respectively, and the cumulative variance explanation rate is 90.170%.

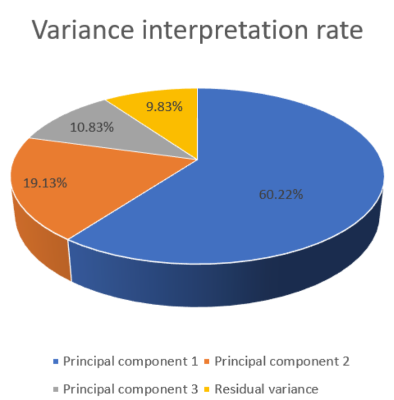


Figure 4.1.2.1 Variance Interpretation Rate

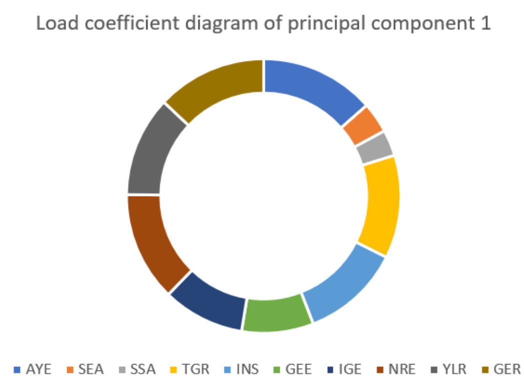


Figure 4.1.2.2 Load Coefficient Diagram

It can be seen from the three-dimensional pie chart 4.1.2.1 that the accumulated three principal components can sum up 90.170% of the information of 10 indicators. Therefore, we can approximate that these three principal components can completely express 10 indicators. The principal component analysis method is successful.

In addition, principal component analysis[3] method can also calculate comprehensive competitiveness.

Table 4.1.2.1 Component Score Coefficient Matrix

| **Component Score Coefficient Matrix** | | | |
|---|---|---|---|
| **Name** | *M* | | |
| | $M_1$ | $M_2$ | $M_3$ |
| *AYE* | 0.164 | -0.061 | -0.025 |
| *SEA* | -0.044 | 0.383 | 0.544 |
| *SSA* | 0.037 | 0.323 | -0.644 |
| *TGR* | 0.148 | -0.192 | -0.036 |
| *INS* | 0.142 | 0.223 | 0.062 |
| *GEE* | 0.103 | 0.396 | 0.055 |
| *IGE* | 0.117 | -0.035 | 0.314 |
| *NRE* | 0.156 | -0.146 | 0.147 |
| *YLR* | 0.143 | 0.026 | -0.26 |
| *GER* | 0.157 | -0.038 | 0.128 |

From Table 4.1.2.1, the relationship between the three principal components M1, M2, M3 and the effective 10 indicators can be obtained. In order to calculate the comprehensive competitiveness, weight the variance interpretation rate of the three principal components.

For example, $M_1$ can be expressed as:

$$\begin{aligned} M_1 &= 0.164\times AYE - 0.044\times SEA + 0.037\times SSA + 0.148\times TGR + 0.142\times INS \\ &+ 0.103\times GEE + 0.117\times IGE + 0.156\times NRE + 0.157\times GER \end{aligned} \quad (19)$$

Final formula $X = 0.668\times M_1 + 0.212\times M_2 + 0.120\times M_3$ (20)

### 4.1.3 BP & PSO Neural Network Builds Health Model

This model uses an evaluation method based on the classification of the higher education system to divide the health evaluation of a country’s higher education into five levels, S, A, B, C, and D, corresponding to scores of 5 to 1, where the rating S represents the country’s The higher education system is very healthy. It is at the top level among all countries. From then on, the letters are progressively worse. The final rating D represents that the country's higher education system is unhealthy compared to other countries, and there are many areas for improvement. In addition, a rating with the symbol **+** indicates that the country's evaluation score is close to the requirements of a higher level. It is worth noting that this article does not use a rating system with the symbol **-**.

In view of the limited number of countries in the world, the amount of data available for training is difficult to meet the needs of the model. These data alone cannot enable the neural network model to obtain better evaluation and prediction capabilities through training. Therefore, after the principal component analysis data is completed, it is necessary to extract a suitable number of principal component factors, and subdivide the factors relative to their upper and lower bounds. At the same time, the subdivided factor sets are graded in stages, so that a large amount of data can be provided as the training set of the neural network. The rating basis after classification is shown in Table 4.1.3.1.

Table 4.1.3.1 Rating Basis

| *Principal component factor* / *Rating* | $M_1$ | $M_2$ | $M_3$ |
|---|---|---|---|
| *S* | $M_1 \geq 0.31$ | $M_2 \geq 0.61$ | $M_3 \geq 0.81$ |
| *A* | $-0.29 \leq M_1 < 0.31$ | $0.01 \leq M_2 < 0.61$ | $0.21 \leq M_3 < 0.81$ |
| *B* | $-0.89 \leq M_1 < 0.29$ | $-0.59 \leq M_2 < 0.01$ | $-0.39 \leq M_3 < 0.21$ |
| *C* | $-1.49 \leq M_1 < -0.89$ | $-1.19 \leq M_2 < -0.59$ | $-0.99 \leq M_3 < 0.39$ |
| *D* | $M_1 < -1.49$ | $M_2 < -1.19$ | $M_3 < -0.99$ |

The training set divided by this rating basis can get better calculation results in neural network training.

In the Back Propagation Neural Network model, this article uses MATLAB to optimize the neural network particle swarm algorithm. In the initialization condition, the algorithm is required to evolve 80 times, and the initial particle swarm number and dimension are 25 and 2, respectively, and the Ackley function is used as the fitness function.When the Back Propagation Neural Network training obtains the error as the fitness value, find and update the corresponding individual and group extreme values, and then update the position and speed. This is repeated until the evolution is completed, and then the optimal weight in this round of solution can be obtained and updated value and threshold, and then the whole neural network is iterated repeatedly until the end. From Figure 4.1.3.1, Figure 4.1.3.2, Figure 4.1.3.4, it can be seen that the use of particle swarm optimization makes the model get faster convergence speed and better solution ability.

After training, the model is used to evaluate the health level of higher education in the two countries in the test set. Of these two countries, one is set as the lower limit of the health level of higher education, and the other is the upper limit. It can be seen from Figure 4.1.3.3 that the education rating obtained after the evaluation of the model can basically fit the pre-set rating. In addition, the accuracy of the test set is calculated to be about 98%, indicating that the model can better analyze the health of higher education in various countries.

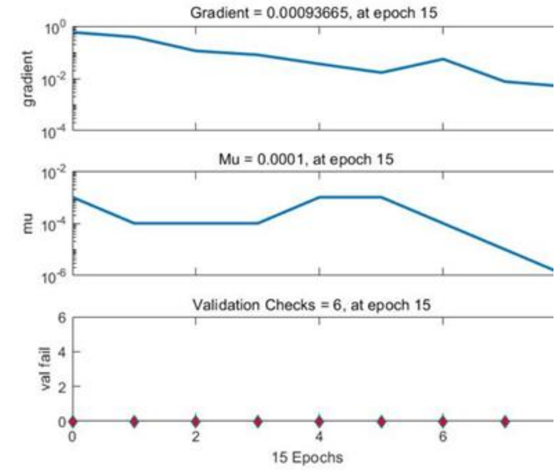


Figure 4.1.3.1 Training State

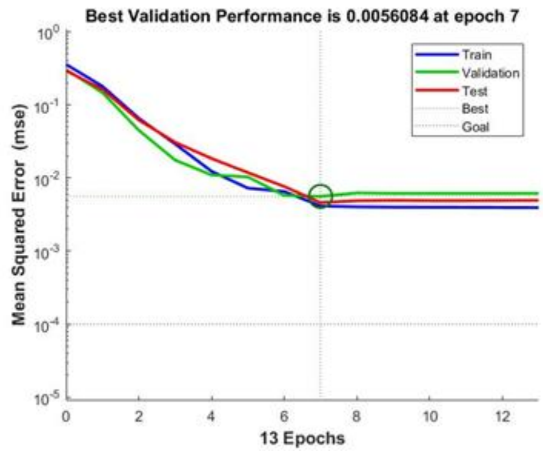


Figure 4.1.3.2 Validation Performance

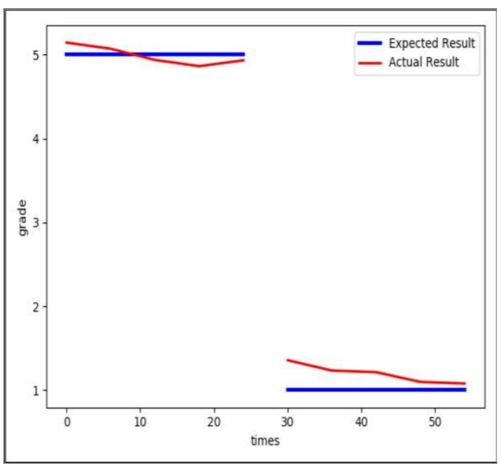


Figure 4.1.3.3 Test-set Performance

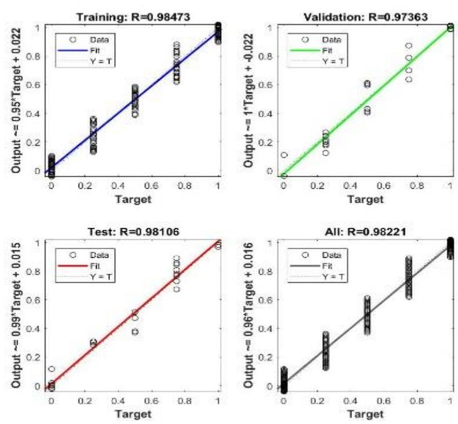


Figure 4.1.3.4 Regression

Finally, the model analyzes the health level of higher education in major countries around the world, and the rating results are shown in Table 4.1.3.2.

Table 4.1.3.2 Country and Score

| **Country** | **Evaluation score** | **Rating** |
|---|---|---|
| Australia | 4.056 | *A* |
| Brazil | 2.803 | *C+* |
| France | 3.220 | *B* |
| United States | 5.034 | *S* |
| Turkey | 1.872 | *D+* |

#### 4.1.4 Grey-ARIMA Hybrid Prediction Method to Establish a Sustainable Model

Based on the current data of the higher education system of various countries, we use the Grey-ARIMA Hybrid Prediction model to obtain future data, and use the BP Neural Network to obtain the health level of the higher education system of each country in the future. Comparing the current health levels with the future health levels of various countries can give a strong indication of sustainability. According to the degree of change of the health level, due to the change degree being higher or lower, the author first forward processes the difference, and then re-trains the neural network, finally the sustainable model of higher education is established precisely. The rating S represents that the country’s high-level sustainability is very good, and the rating D represents that the country’s high-level sustainability is relatively poor.

After the PCA method, there are no longer needs to separately predict the ten indicators of the five countries, which will cause a large amount of calculation. The author only needs the three principal components of the five countries to Grey- ARIMA Hybrid Prediction method. The prediction of 15 indicators is compared with the 50 indicators before dimensionality reduction, which shows that the PCA method is very effective. The following takes Turkey’s principal component three prediction as an example to illustrate the process of the prediction.

The first is grey prediction. The GM(1,1) model is constructed. After passing the SPSS analysis, first perform the grade ratio test. The result of the grade ratio test of the data column is 0.787, which is within the applicable range [0.717, 1.396], indicating that the data column is suitable for the GM(1,1) model, which is gray The prediction can get a satisfactory model.

Table 4.1.4.1 Model Construction Results

| Development coefficient a | Grey action b | Posterior difference ratio C |
|---|---|---|
| -0.2302 | -0.2323 | 0.0999 |

Development coefficient a and Grey effect b are the output values of model construction; Posterior difference ratio C is used for the model accuracy level test, C<0.35 can indicate the model accuracy level is good, C<0.65 indicates the model accuracy is good, but ,When C>0.65, it fails the model accuracy test and is not suitable for grey prediction. The posterior difference ratio of this paper is C=0.100<=0.35. According to the test rules, the accuracy of this model is good, which is more suitable for gray prediction and can get more accurate prediction values.

After the model is constructed, the relative error and the level ratio deviation can be analyzed to verify the effect of the model.

Table 4.1.4.2 GM(1,1) Model Test

| Serial number | Original value | Predictive value | Residual | Relative error | Step ratio deviation |
|---|---|---|---|---|---|
| 1 | -1.581 | -1.581 | 0 | 0.00% | 0 |
| 2 | -1.356 | -1.351 | -0.005 | 0.34% | -0.061 |
| 3 | -1.232 | -1.23 | -0.002 | 0.18% | 0.001 |
| 4 | -1.103 | -1.119 | 0.016 | 1.43% | -0.016 |
| 5 | -1.03 | -1.018 | -0.012 | 1.16% | 0.026 |

It can be seen from the above table that the Residuals, Relative errors and Steps ratio deviations of the predicted values corresponding to the five original values all meet the minimum requirements, which can indicate that the gray prediction model has a better fitting effect on the future principal component data of Turkey.

Next is ARIMA prediction. Using the principle of minimum AIC value, through SPSS to find various possible model combinations for model construction, and obtains the most suitable AR model, difference order and MA model, and finds the autoregressive order p, the difference order Number d value and moving average order q.

Table 4.1.4.3 MA(1) Model Parameter

| Item | Symbol | Value |
|---|---|---|
| Constant term | c | -1.281 |
| MA parameters | β1 | 1.000 |
| Q statistics | Q6(*p*) | 0.021(0.886) |
| Information Guidelines | AIC | 1.542 |
| | BIC | 0.37 |

* $p$<0.05 ** $p$<0.01

The above table is the result of the model construction of the principal component three of Turkey, and the model formula is:

$$y(t) = -1.281 + 1.000 \times \varepsilon(t-1) \tag{21}$$

From the Q statistic results, the p>0.1 of Q6 means that the null hypothesis cannot be rejected at the significance level of 0.1, and the null hypothesis is established. The residual of the ARIMA model is white noise, and the model basically meets the requirements.

The following are two figures showing the prediction and fitting of the principal component three of Turkey by the gray prediction model and the ARIMA model.

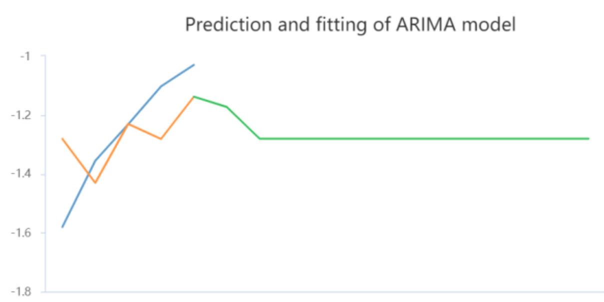


Figure 4.1.4.1 ARIMA Model

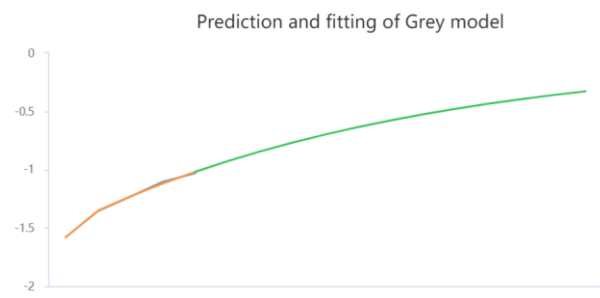


Figure 4.1.4.2 Grey Model

It can be seen that both the Grey Prediction Model and the ARIMA Model can fit the original data for the next 12 periods, but there are certain differences in the fit of the two models. That is to say, the forecast data for the first to third periods after the original data are more accurate results, and the prediction data for the subsequent 4-12 periods are not well fitted.

The author averages the ten indicators of all five countries obtained by the gray forecast model and the ARIMA model, and obtains the following forecast data, which is used to further build a sustainable model of the higher education system.

Table 4.1.4.4 The Average Prediction Value

| Country | $M_1$ | $M_2$ | $M_3$ |
|---|---|---|---|
| Australia | 0.839 | -1.3735 | 1.1215 |
| Brazil | -0.7575 | 0.667 | 0.4635 |
| France | 0.4195 | -0.4495 | -1.9715 |
| United States | 1.3575 | 1.914 | 0.211 |
| Turkey | -1.024 | 0.264 | -0.7415 |

The above table is the result of the Grey-ARIMA Hybrid Prediction model of the three principal components of five countries and the average value.

Make the difference between the predicted data of the three principal components and the pr evious three principal components.Since the difference data has a positive and negative relationship, the difference between the forecast data and the current data should be forward processed, and substituted it into the Neural Network to retrain. From this, we can get an explanation of the sustainability rating. The rating S represents the sustainability of the country's higher education system is very good, and the rating D represents the sustainability of the country's higher education system is relatively poor. The description of the rating process of the neural network training set is omitted here, which is similar to 4.1.3 above.

The rating results of the sustainability of higher education in major countries in the world are directly obtained as shown in the following table.

Table 4.1.4.5 Country and Evaluation Score

| **Country** | **Evaluation score** | **Rating** |
|---|---|---|
| Australia | 1.743 | *B+* |
| Brazil | 0.217 | *D+* |
| France | 1.409 | *B* |
| United States | 2.649 | *S* |
| Turkey | 0.883 | *C+* |

## 4.2 Application and Analysis of Higher Education Model

After passing the verification in five countries, in order to apply the health model and sustainability model of the higher education system we established to a wider range of countries, this paper added new data from Japan and Germany to the model, the health level and and sustainable level are shown in Table 4.2.1.

Table 4.2.1 New Country and Score

| **Country** | **Health score** | **Health rating** | **Sustainability score** | **Sustainability rating** |
|---|---|---|---|---|
| Japan | 4.807 | *A+* | 2.416 | *S* |
| Germany | 4.629 | *A+* | 1.992 | *A* |

After adding data from Japan and Germany, the health and sustainability ratings of eight countries were compared more clearly and intuitively. (Including the United States, Japan, Germany, Australia, France, Brazil and Turkey) The author made a three-dimensional tree diagram, as shown in the figure below:

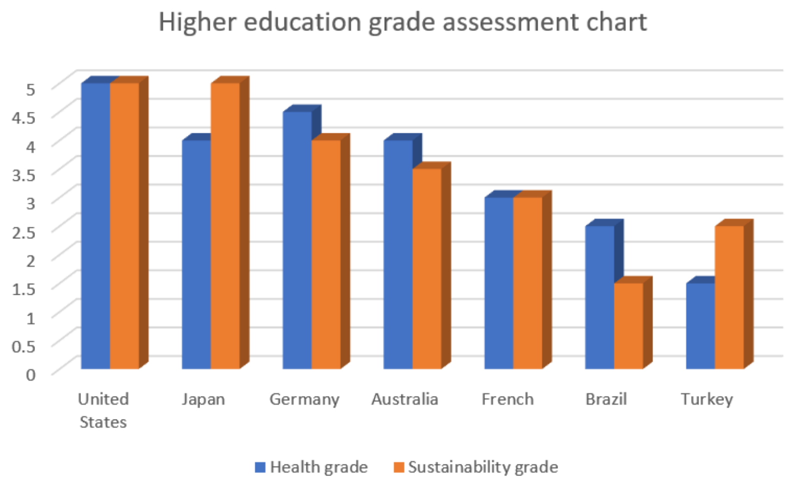


Figure 4.2.1 Grade Assessment

From the three-dimensional tree diagram 4.2.1, it can be seen that the United States has the highest level of S in both the health rating and the sustainability rating, indicating that the United States is ahead of other countries in terms of higher education convenience. Japan ranks second overall. Japan’s health rating is slightly worse than the United States, but the sustainability rating is the same as the United States. In the future, whether the United States can continue to maintain its leading position in the level of higher education still depends on whether future policies can effectively increase the values of various indicators. After that, this paper chose Brazil as an example for political planning and model prediction of higher education promotion.

## 4.3 A Higher Education Promotion Program Taking Brazil as an Example

### 4.3.1 Brazil's Health and Sustainability Analysis

The health rating of Brazil's higher education system in the world is C+, and the sustainability rating is D +, which is generally poor compared with other countries. It is mainly limited by the following factors, which need to be improved in the future.

1.Brazil's years of education per capita are relatively short, which indicates that some people have not got enough education time, so that they can not enter higher education, thus reducing the talent reserve capacity of higher education.

2.The number of Brazilian students studying abroad is several times the number of students entering the country, which will cause the loss of Brazil's outstanding talents and lead to the decline of Brazil's higher education system.

3.The government’s per capita expenditure on higher education is relatively low compared to other countries, indicating that Brazilian students receive insufficient subsidies and benefits in higher education, which is not conducive to providing students with sufficient development opportunities.

4.The government's scientific research expenditure is also relatively small compared to other countries, which is not conducive to the development of scientific research in higher education, which also affects many evaluation indicators.

### 4.3.2 Targeted Policies and Timetable

In response to the weaknesses of Brazil’s current education system, the designated schedule is as follows.

Table 4.3.2.1 Policy Timeline

| Index | The first year | The second year | The third year |
|---|---|---|---|
| *NRE* | 1.34% | 1.68% | 2.34% |
| *PEE* | 29% | 30% | 32% |
| *AYE* | 7.8 | 7.9 | 8.0 |
| *GER* | 48% | 50% | 53% |

The principal component analysis of the third year's required indicators will find that the principal component factors have significant changes. Then, substituting the factors into the trained neural network model will find that Brazil’s higher education health rating has changed from C+ to B. The sustainability rating has changed from D+ to C+, which indicates that these policies are very effective and can improve the level of Brazil's higher education.

## 4.4 Realistic Impact of Higher Education

Throughout the ages, no matter which aspect it is aimed at, the implementation of a new policy cannot be easy, and it is often accompanied by many challenges and difficulties. Brazil’s economic development has slowed down in recent years, so it needs to make great efforts to increase research expenditures and higher education expenditures. At the same time, extending Years of education per capita and enrollment rate of higher education means that not only a large amount of investment is required, but also the need to ensure that the policy can be implemented for a long time. During the transition period, the implementation of new policies will inevitably encounter resistance due to changes in government decision-making or changes in the world economic situation. Some policies even fail because they are difficult to implement. Therefore, the success of any change takes time to prove.

When a country's higher education system has undergone significant changes, first of all, from the national level, the country should actively introduce relevant policies in conjunction with the new system to promote the implementation of the new system. secondly, from the school level, the new system, the new policy also means a new beginning. Schools should implement relevant policies issued by the state, and carry out the reform in combination with their own school running characteristics. Teachers, schools and students are closely related and are an inseparable whole. Teachers, schools, and students are closely related and are an inseparable whole. Teachers should actively cooperate with the school's new policy, correctly guide students, and fully implement the new policy step by step. The implementation of the new policy faces challenges at all levels. If all levels cooperate in an orderly manner, the new policy will bring unexpected results; however, if one party has errors in the implementation process, the final result will be may not be so perfect.

# 5. Evaluation, Improvement and Extension of the Model

## 5.1 Evaluation of the Model

**Advantages:**

1.This set of models has strong applicability and is suitable for evaluation and prediction problems in many different fields. It is more classic.

2.In the process of data processing, methods such as averaging, normalization, and standardization are used to make the solution and verification of the model more accurate.

3.Using the principal component analysis to reduce the dimensions of the indicators, the correlation between the evaluation indicators is eliminated, and the average training time and average number of iterations of the Back Propagation Neural Network are reduced.

4.Using Back Propagation Neural Network to classify higher education is more objective than fuzzy evaluation method and analytic hierarchy process.

5.Use the PSO algorithm to optimize the Back Propagation Neural Network, which improves the convergence speed and solution accuracy.

6.Reasonably use Grey-ARIMA hybrid forecasting method, establish a sustainable model, and effectively reduce the error of the forecast value.

**Disadvantages:**

1.The 12 indicators established by the participation model cannot fully measure the level of the higher education system under actual conditions.

2.There are still some subjective factors in the classification of Back Propagation Neural Network.

3.The adjustment of different structures and parameters will make the PSO algorithm have better room for improvement in convergence speed and efficiency.

4.There are a lot of data involved in the model, which puts forward some requirements on the amount of calculation.

5.In the process of data collection, the lack of the most recent 2020 data will have a certain impact on the forecast results.

## 5.2 Improvement of the Model

**Adjustability of Neural Network Parameters:** In order to get better learning ability and faster convergence speed, the model can be improved by learning rate attenuation, modifying the number of nodes in the hidden layer, using momentum gradient descent[4] and so on.

1. Decline of learning rate

The learning rate lr is between [0-1]. During the training process, if the value of lr is relatively small, the convergence speed during the training process will decrease. If the value of lr is relatively large, intense oscillations may occur during the training process, resulting in a decrease in the accuracy of the best model. The learning rate can be accelerated by learning rate decay, that is, lr gradually decreases during the training process.

$$lr = lr0/(1+dr\times p) \tag{22}$$

(22)lr0 is the initial learning rate, dr is the decay rate, and p is the training batch.

2. Modify the number of hidden layer nodes

In this model, the accuracy of the model can be tested by gradually testing the number of nodes in the hidden layer. In fact, there is an empirical formula that can roughly predict the appropriate number of hidden layer nodes, the formula is as follows.

$$m = \sqrt{n+1}+\alpha \tag{23}$$

$$m = \log_2 n \tag{24}$$

$$m = \sqrt{nl} \tag{25}$$

(22)-(25)m is the number of nodes in the hidden layer, n is the number of nodes in the input layer, l is the number of nodes in the output layer, α is a constant, and its value range is [1,10]. In this model, the correct rate of different hidden layer nodes is shown in Figure 5.2.1.1.

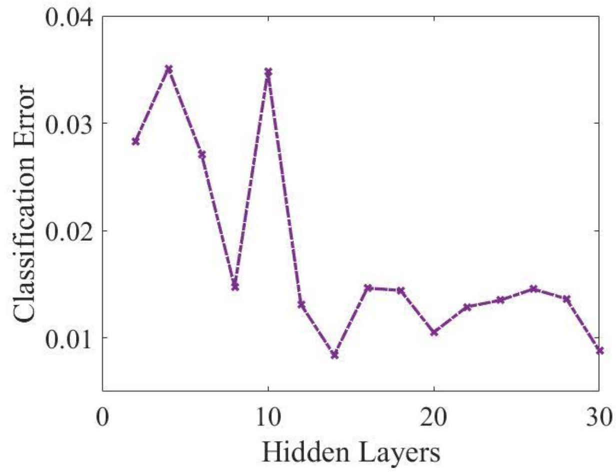


Figure 5.2.1 Hidden Layers

3. Momentum gradient descent

Back Propagation Neural Network essentially uses the gradient descent, which may fall into local minimum without additional momentum. Therefore, a momentum is needed to make it out of the range of the local minimum, which can significantly improve the accuracy of the model. The formula of the momentum gradient descent method is as follows.

$$W(k+1) = W(k) + lr[(1-n)D(k) + nD(k-1)] \tag{26}$$

(26)W is the weight, D(k) and D(k-1) represent negative gradients of K and k-1 degree, lr is the learning rate, and n is the momentum value between [0-1].

## 5.3 Extension of the Model

The model in this article is a typical evaluation and prediction model. In addition to the field of higher education, it can also be applied to the evaluation and prediction of the ecosystem and the evaluation and prediction of the economic situation. In addition to principal component analysis and Back Propagation Neural Network, factor analysis, fuzzy evaluation, multiple regression model can also be used to replace in a more suitable data set ,so that the model can be applied to a wider range of scenarios.

# 6. Sensitivity Test

Sensitivity analysis[5] is of great significance in this higher education evaluation model. Because through the sensitivity analysis, we can know which indicators will have a greater impact on the model, so we can know which indicators in the higher education system should be as stable as possible, so we can put forward some more reasonable suggestions for the country.

First, this paper tests the global sensitivity analysis test to verify whether the higher education evaluation model is stable. The Monte Carlo algorithm is used to randomly perturb 12 indexes, and the deviation percentages are 2% and 5% respectively. The comprehensive score is obtained by neural network after the perturbed data, and the error rates are 6.37% and 10.26% compared with the previous comprehensive score, which pass the sensitivity test. Next, this paper conducts local sensitivity analysis. Because local sensitivity analysis of each indicator in the model requires a large amount of calculation, Therefore, according to the score coefficient matrix of each index in the principal component analysis, the author selects GEE and TGR for analysis, and finds out the sensitivity S is 3.5% and 7.2%. It can be seen that the output comprehensive score of education level has little change, and the sensitivity of the model is better. At the same time, it can be seen that TGR has a greater impact on higher education than GEE.

# References

[1] Zuo Chongliang, Pan Maoyuan. The core principles and internal and external patterns of American h igher education governance[J]. Jiangsu Higher Education, 2016(06): 24-30.

[2] Luo Hongben. Research on Intelligent Hybrid Forecasting of Financial Time Series Based on Grey-ARIMA[J]. Finance and Economics Theory and Practice, 2014, 35(02): 27-34.

[3] Nie Min, Liu Zhihui, Liu Yang, Yao Junqiang. Runoff prediction based on PCA and BP neural netw ork [J]. China Desert, 2016, 36(04): 1144-1152.

[4] Peng Xianlun. Convergence of momentum gradient descent method [D]. East China University of sci ence and technology, 2020

[5] Xu Chonggang, Hu Yuanman, Chang Yu, Jiang Yan, Li Xiuzhen, Bu Rencang, He Hongshi. Sensitiv ity analysis of ecological model[J]. Journal of Applied Ecology, 2004(06):1056-1062.

# Appendix

1. Part of the original data used in this paper.

| time | years of education aged 25 ar | s in all fields of each country (fractional mea | higher education students stud | f bachelor's degree in higher educ | of international s | diture on education, calculated at purchasing power parity |
|---|---|---|---|---|---|---|
| Australia | 12.3 | 21,015 | 12372 | 72.40121 | 266048 | 15084.73042 |
| 5 | 12.5 | 21,700 | 12143 | 58.29639 | 294438 | 16890.83534 |
| 6 | 12.6 | 22,253 | 12849 | 60 | 335512 | 16213.3902 |
| 7 | 12.7 | 22,543 | 13558 | 62.17685 | 381202 | 15483.58194 |
| 8 | 12.7 | 22,405 | 13319 | 65.31132 | 444514 | 15600 |
| Brazil | 7.4 | 62,104 | 45270 | 21.45059 | 19093 | 36541.72058 |
| | 7.6 | 63,550 | 50808 | 22.07671 | 19855 | 40366.26016 |
| | 7.7 | 66,813 | 52513 | 26.234 | 19996 | 40703.36887 |
| | 7.8 | 70,115 | 58876 | 27.49514 | 20671 | 45884.88978 |
| | 7.8 | 73,073 | 67183 | 48.32423 | 21181 | 46000 |
| France | 11.4 | 106,435 | 82073 | 49.85922 | 235123 | 35000.20938 |
| | 11.5 | 106,099 | 86684 | 53.28056 | 239409 | 35869.50206 |
| | 11.4 | 106,846 | 90842 | 52.8997 | 245349 | 36759.22792 |
| | 11.4 | 107,448 | 89412 | 55.0193 | 258380 | 38845.37419 |
| | 11.4 | 104,354 | 99488 | 58.47153 | 229623 | 39100 |
| United State | 13.3 | 535,617 | 77963 | 60.10556 | 842384 | 229044.6404 |
| | 13.3 | 537,423 | 80560 | 58.48336 | 907251 | 239344.4556 |
| | 13.3 | 541,080 | 83984 | 54.15031 | 971417.0195 | 220602.7359 |
| | 13.4 | 552,148 | 86596 | 53.55469 | 984897.6084 | 272161.3438 |
| | 13.4 | 548,847 | 84349 | 53.08304 | 987313.6528 | 290000 |
| Turkey | 7.7 | 35,967 | 45189 | 26.31448 | 48183 | 24364.27677 |
| | 7.6 | 38,102 | 45684 | 28.58708 | 72178 | 25051.11404 |
| | 7.8 | 41,005 | 45545 | 28.43483 | 87903 | 24665.6135 |
| | 8 | 39,442 | 45736 | 27.91645 | 108076 | 22507.1621 |
| | 8.1 | 39,449 | 47546 | 27.65687 | 125138 | 23000 |

| Youth literacy rate | Global higher education enrollment rate (% of total) | rity Index for Higher Educatio | ment expenditure for higher education students (% d | ational R&D expenditure (% of GD |
|---|---|---|---|---|
| 99.41957092 | 79.15869904 | 1.3 | 37.06478 | 3.08429 |
| 99.4389801 | 80.71942139 | 1.307700038 | 36.28273 | 3.04979 |
| 99.45876312 | 83.45494843 | 1.300559998 | 35.81112 | 3.12588 |
| 99.47612 | 85.05713654 | 1.296229959 | 35.98491 | 3.05212 |
| 99.4960022 | 86.68834686 | 1.262030005 | 36.23652 | 3.17177 |
| 98.84568787 | 43 | 1.220260024 | 27.41477 | 1.19567 |
| 98.96375275 | 44 | 1.227830052 | 27.88126 | 1.27131 |
| 99.11096191 | 45 | 1.232730031 | 29.81738 | 1.34264 |
| 99.15541077 | 46 | 1.235000014 | 28.79964 | 1.26417 |
| 99.20417023 | 47 | 1.231649995 | 33.04323 | 1.26326 |
| 99.68022156 | 61.51047897 | 1.195440054 | 34.06808 | 2.27592 |
| 99.67006683 | 62.78593063 | 1.19532001 | 33.75392 | 2.26703 |
| 99.66880798 | 64.72767639 | 1.191840053 | 33.43725 | 2.22238 |
| 99.68 | 65.80223083 | 1.194540024 | 32.96949 | 2.20557 |
| 99.70707703 | 67.62465668 | 1.197710037 | 31.62493 | 2.20002 |
| 99.71957092 | 88.6268692 | 1.265130043 | 21.89296 | 2.71924 |
| 99.7389801 | 88.88941193 | 1.26267004 | 20.41236 | 2.71742 |
| 99.75876312 | 88.83505249 | 1.25643003 | 21.16585 | 2.76145 |
| 99.77612 | 88.16738892 | 1.26481998 | 21.52057 | 2.81741 |
| 99.7960022 | 88.29917908 | 1.268270016 | 19.42953 | 2.83766 |
| 98.22458649 | 48.85942078 | 0.867370009 | 32.71303 | 0.81821 |
| 98.70935059 | 50.76102829 | 0.873369992 | 31.19523 | 0.86077 |
| 99.01602173 | 51.84365082 | 0.872049987 | 31.48285 | 0.8815 |
| 99.22743225 | 52.69842148 | 0.879809976 | 30.269 | 0.94464 |
| 99.35848236 | 52.6969986 | 0.898270011 | 32.28416 | 0.96105 |

2. Example of the code used in this paper.

```python
from sklearn.neural_network import MLPRegressor
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

data = pd.read_excel("./training.xlsx")
test = pd.read_excel("./testing.xlsx")
fin = pd.read_excel("./MainIngred.xls")
pred = pd.read_excel("./predict.xlsx")

model = MLPRegressor(hidden_layer_sizes=(4,), random_state=0.01, learning_rate_init=0.01)
zi = data.iloc[:, :3]
yi = data.iloc[:, 3]
model.fit(zi, yi)
predicts = []
for line in pred.index:
    line = pred.loc[line].values[0:3]
    res = model.predict([line])
    predicts.append(res)
    print(res)

xx = range(0, int(len(predicts) / 2))
xx2 = range(int(len(predicts) / 2), len(predicts))

plt.plot(xx, test.iloc[:, 3][0:5], color="blue", label="Expected Result", linewidth=3)
plt.plot(xx2, test.iloc[:, 3][5:], color="blue", linewidth=3)
plt.plot(xx, predicts[0:5], color="red", label="Actual Result", linewidth=2)
plt.plot(xx2, predicts[5:], color="red", linewidth=2)
plt.xlabel("times")
plt.ylabel("grade")
plt.legend()
plt.show()
```