Advanced Science and Technology Letters
Vol.111 (NGCIT 2015), pp.179-184
http://dx.doi.org/10.14257/astl.2015.111.35

# Prediction of Educational Investment in China

[1]Xiao Qianyin, [2]Liu Bo

[1]Foreign Language School, Southwest Petroleum University, Chengdu Sichuan, 610500, China, E-mail: conniexqy@qq.com
[2]Department of Planning and Evaluation (Teacher Education and Development Center), Southwest Petroleum University, Chengdu Sichuan, 610500, China, E-mail: th100th@swpu.edu.cn

**Abstract.** In this article, we aim at presenting a series of solutions for the prediction of educational investment for China based on multiple linear regression (MLR), artificial neural networks (ANNs) and grey model GM (1,1). Multiple comparisons are made for deciding whether model should be used under different external conditions. Our research successfully shows that all these models are available for practical applications and scientists and other related people can choose their suitable models alternatively for the sake of making a better prediction under different circumstances.

**Keywords:** Educational investment; Chinese education; multiple linear regression; grey model GM (1,1); artificial neural networks

## 1 Introduction

In this paper, we aim at presenting a series novel method for the prediction of educational investment for China. According to the complex linear or non-linear relationship of social indicators, there are some indicators that are commonly considered as the impact factors of educational investments in China. These indicators consist of a series of economical conditions and development of China, including gross domestic product, the number of students in ordinary colleges and Universities, social organizations and citizens to education, as well as social donations and fund-raising education. All these four respects are considered as crucial indicators according to previous related research. Here, we use these four indicators as the independent variables for the development of relevant linear and non-linear models. And the educational investment is set as the dependent variable. One can only input the data of independent variables to obtain the dependent variable in the computer, using the linear and non-linear models presented by this article.

ISSN: 2287-1233 ASTL

## 2 Artificial Neural Network

Artificial neural network (ANN) is an information processing system with interconnected components analogous to neurons [5-10]. The system has splendid nonlinear mapping and a good degree of parallel processing of information.

## 3 Grey Model GM (1, 1)

Grey Model GM (1,1) is an excellent mathematical approach that is mainly used for predicting in uncertain systems. Original data can be changed into new forms for the sake of finding out regulations of the change of the system, which is mainly determined by a series of iteration. Prediction based on GM (1,1) is thoroughly independent to the independent variables. GM (1,1) is widely used for the prediction of the values which change with a certain gap of time.

## 4 Results and Discussion

In this section, the MLR, ANN models the GM (1,1) model are presented for the prediction of educational investments respectively, based on the data provided by reference [18]. Results and establishing processes of all the three kinds of models are presented in the separated sections.

### 4.1 Development of the MLR Model

Here, we use SPSS to develop an equation for the change regulation of educational investments in China. The 4 indicators were presented by $x_1$ to $x_4$, whereas $y$ is the educational investments. The multiple linear regression equation (Equation (1)) is shown as follows

$$y = 0.047x_1 + 3.781x_2 - 7.257x_3 - 2.914x_4 - 1325.869 \quad (1)$$

Equation (1) is the result of MLR modeling, based on the existing data of educational investments as well as the four indicators, which shows that $x_2$, $x_3$, $x_4$, have significant impacts to the educational investment of China ( which presented as $y$ in the equation). Fig. 1 illustrates the comparison of actual values versus the predicted values using the multiple linear regression method:

 

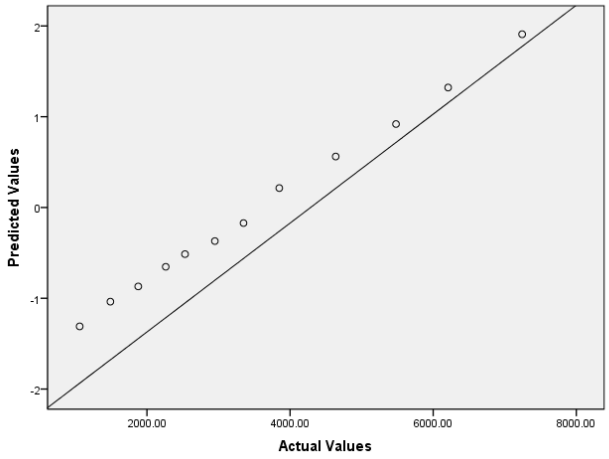


**Fig. 1.** Actual vales versus predicted values via the MLR model.

Fig. 1 illustrates that the MLR model has a decent result during the prediction. However, we still consider that the results are not precise enough because the linear function cannot directly fit the relationship between the educational investments and the four indicators. Therefore, we decided to use non-linear functions to fit the data. An artificial neural network is a powerful machine learning techniques that can fit the data into a non-linear form. In Section 4.2, we present the detailed results of the training and testing processes of relevant ANN models.

### 4.2 Development of ANN Models

For the sake of developing a series of ANN models, we used the same indicators and dependent variable. Best net search method was used for finding a suitable ANN model. We used the general regression neural network (GRNN) [19-23] and multi-layer feed-forward neural network (MLFN) [24-27] to develop the models and undertake the best net search comparison (see Table 1). The numbers of nodes of MLFN models were set from 2 to 25 in order to find out the most suitable nodes of MLFN model for the prediction.

Table 1 shows the best net search results of ANN models for the prediction of the emission of HC in the first system:

**Table 1.** Best net search results of models for the prediction of private vehicle ownership in Chinese area.

| ANN Model | Trained Samples | Tested Samples | RMS Error | Training Time |
|---|---|---|---|---|
| GRNN | 8 | 4 | 254.56 | 0:00:00 |
| MLFN with 2 Nodes | 8 | 4 | 462.55 | 0:00:34 |
| MLFN with 3 Nodes | 8 | 4 | 1121.79 | 0:01:05 |

| | | | | |
|---|---|---|---|---|
| MLFN with 4 Nodes | 8 | 4 | 1242.45 | 0:01:42 |
| MLFN with 5 Nodes | 8 | 4 | 2887.77 | 0:04:54 |
| MLFN with 6 Nodes | 8 | 5 | 48437.35 | 0:26:16 |
| MLFN with 7 Nodes | 8 | 4 | 438.72 | 0:34:28 |
| MLFN with 8 Nodes | 8 | 4 | 7116.37 | 0:00:31 |
| MLFN with 9 Nodes | 8 | 4 | 9775.28 | 0:00:33 |
| MLFN with 10 Nodes | 8 | 4 | 1390.46 | 0:00:38 |
| MLFN with 11 Nodes | 8 | 4 | 4370.56 | 0:00:39 |
| MLFN with 12 Nodes | 8 | 4 | 3259.50 | 0:00:47 |
| MLFN with 13 Nodes | 8 | 4 | 6869.84 | 0:00:46 |
| MLFN with 14 Nodes | 8 | 4 | 4927.99 | 0:00:48 |
| MLFN with 15 Nodes | 8 | 4 | 5121.20 | 0:00:53 |
| MLFN with 16 Nodes | 8 | 4 | 7709.36 | 0:00:58 |
| MLFN with 17 Nodes | 8 | 4 | 17625.61 | 0:01:03 |
| MLFN with 18 Nodes | 8 | 4 | 6990.54 | 0:01:06 |
| MLFN with 19 Nodes | 8 | 4 | 3216.10 | 0:01:04 |
| MLFN with 20 Nodes | 8 | 4 | 11387.71 | 0:01:05 |
| MLFN with 21 Nodes | 8 | 4 | 12657.03 | 0:01:12 |
| MLFN with 22 Nodes | 8 | 4 | 2720.15 | 0:01:18 |
| MLFN with 23 Nodes | 8 | 4 | 10426.98 | 0:01:23 |
| MLFN with 24 Nodes | 8 | 4 | 7360.01 | 0:01:29 |
| MLFN with 25 Nodes | 8 | 4 | 12855.64 | 0:01:38 |

Table 1 indicates that the GRNN is the best model for prediction, with an lowest RMS error (254.56) in the result list. Figure 3 is illustrating the change regulation of the RMS errors of MLFN models in accordance with the change of nodes:

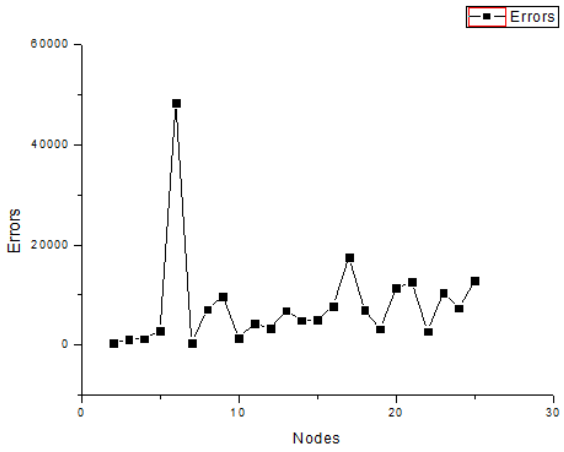


**Fig. 2.** RMS errors versus nodes of MLFN models

Fig. 2 shows that there is a significant fluctuation of RMS errors when changing the nodes of MLFN model. The fluctuation presented by Figure 3 indicates that the MLFN model is not as suitable as the GRNN model. And the overall RMS errors of MLFN models during our experiments are not as low as that of GRNN model. Therefore, we strongly believe that the GRNN model is the best model during the process of best net search. In addition, the results below presents the training and testing process of the GRNN model only because the GRNN model is believed to be the only suitable ANN model during our detailed repeated experiments.

Here, we present the results of the GRNN model in Fig. 3

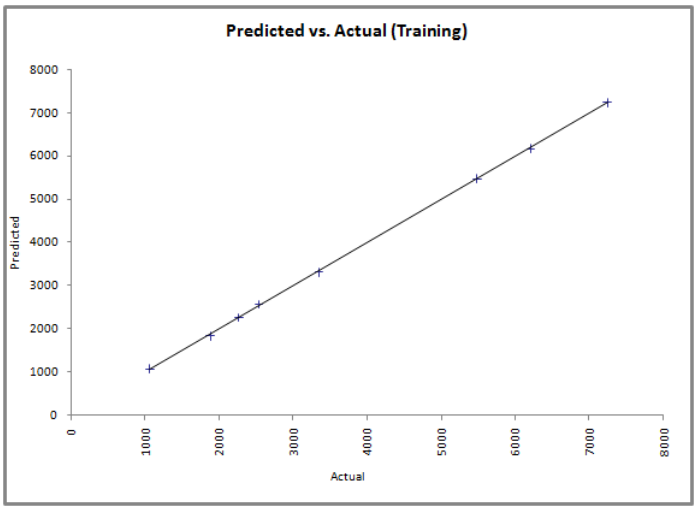


**Fig. 3.** Predicted values versus actual values after training in GRNN

Fig 3 shows that the training process is extremely robust. The regression line presents a regulation of proportional function. Hence, this result shows that the training process is correct and precise. It also confirms that Figs. 5 and 6 is reliable.

### 4.3 Development of Grey Model GM (1,1)

A GM (1,1) model based on the sequence of educational investments was established using relevant software. 8 data was used for the model development, whereas 4 data was used for testing. Testing results are presented by Fig. 4.

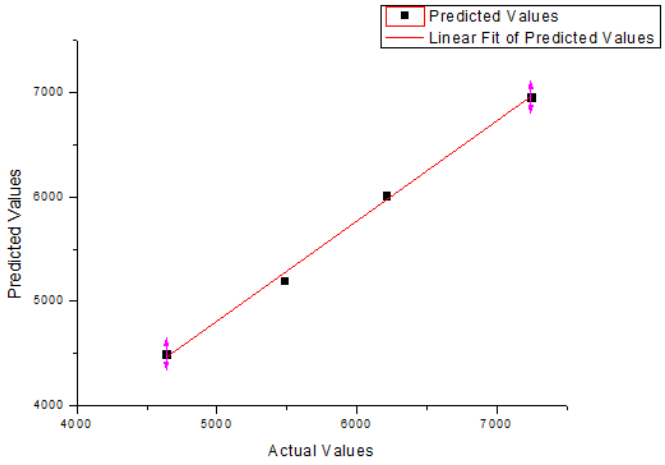


**Fig.4.** Testing result of grey model GM (1,1) (actual values versus predicted values).

Fig. 4 presents the testing results of GM (1,1) model for the prediction of educational investment in China. Results show that the testing situation of GM (1,1) model is nearly as good as that of GRNN model, with a very suitable result. One of the advantages of GM (1,1) model is time-saving, and it can be used for predicting more data in the future years. However it can only be used for the prediction to the following future years because GM (1,1) model is only perfectly suits for a exponential prediction. If the educational investment is not corresponding to the exponential changes, in the long run, the prediction result may fail to correspond to actual results.

## References

1. Y. L, P. L., Economic Research Journal, 3 (2003)
2. J. Monson, The China Quarterly, 200 (2009)
3. G. Montinola, Y. Qian, B. R. Weingast. Federalism, World Politics, 48, 1 (1995)
4. B, Jickling, The Journal of Environmental Education, 23, 4 (1992)
5. T. Gylfason European Economic Review, 45, 4 (2001)
6. B. Samanta, K. R. Al-Balushi. Mech. Syst. Signal Pr., 17, 2 (2003)
7. W. G. Baxt, Ann. Int. Med., 115,11, (1991)
8. M. Shao, X. J. Zhu and H. F. Cao, Energy, 67 (2014)
9. C. H. Aladag, A. Kayabasi and C. Gokceoglu, Neural Comput. App., 23, 2 (2013)
10. J. Van Schependom, G. Nagels and W. Yu, Schizophrenia Bulletin (2013).
11. E. C. Santos, E. D. Armas and D. Crowley, Soil Bio. Bioche., 69 (2014)
12. H. Yip, H. Fan, Y. Chiang, Auto. Const. 38 (2014)
13. C. M. Hong, F. S. Cheng, C. H.Chen, Int. J. Elect. Power & Energy Syst., 60 (2014)
14. M. Mao and E. C. Chirwa. Technological Forecasting and Social Change, 73, 5 (2006)

 