An international bank failure prediction model based on BP neural network

Hongjie Yue

Kharkiv Institute, Hangzhou Normal University, Hangzhou, Zhejiang, 311121, China

Abstract: International bank failures occur frequently. In order to analyze and predict the causes of international bank failures, this paper uses the index data of Polish banks from 2017 to 2021 to establish a BP neural network prediction model. The optimal BP neural network prediction model with 50,000 training times, low time complexity and high model accuracy was determined.

1. Introduction

Banks play an important role in the development of national economy and society. The bankruptcy of banks brings many adverse effects to society, enterprises and individuals[1].Compared with domestic banks, international banks fail frequently, so the analysis and prediction of their failure causes become the core issue of research. Consider relevant data bank of large sample size and a lot of data and the characteristics of high correlation between indicators, and the characteristics of the multiple regression model accuracy is not high, in this paper, neural network prediction model is established, with bank data set of 64 indicators for the input set, with each sample (bank) composite scores as the output of the training, Then the value and weight of each hidden layer of neural network are obtained[2]. Considering the large capacity 64 indicators data sets, using the index of bank failures to data mining, the cause of the first principal component analysis (pca) is used to analyse the sample data dimension reduction, induces the most information can be summarized in the sample data of five main components, five main ingredients after extraction of the covers 61.453% of the information, The comprehensive score index dataset was obtained by using the feature vectors and standardized dataNext, the combined score of each bank is predicted. Finally, the comprehensive score of the prediction is analyzed, and the accuracy of the model is evaluated by relative deviation.

2. Establishment of BP neural network prediction model

2.1 BPNN is briefly

BPNN is the BP neural network, which includes three parts: input layer, hidden layer and output layer[3].Taking the third layer as an example, assuming that the input feature vector is, the output of its hidden layer can be expressed as $x = (x_1, x_2, \cdots, x_n)$

$$y_i = f(\sum_j w_{ij} x_j - \theta_j) \tag{1}$$

Where, is the output of the first neuron in the hidden layer; $y_i i x_j$ Is the first feature; $j w_{ij}$ Is the weight value from the first neuron in the input layer to the first neuron in the hidden layer; $j i \theta_j$ Is the threshold of the first neuron in the hidden layer; $j f(\cdot)$ Is the transfer function from the input layer to the hidden layer. The output of the output layer is similar to the output of the hidden layer[4]. The structure diagram of BP neural network is shown in Figure 1.
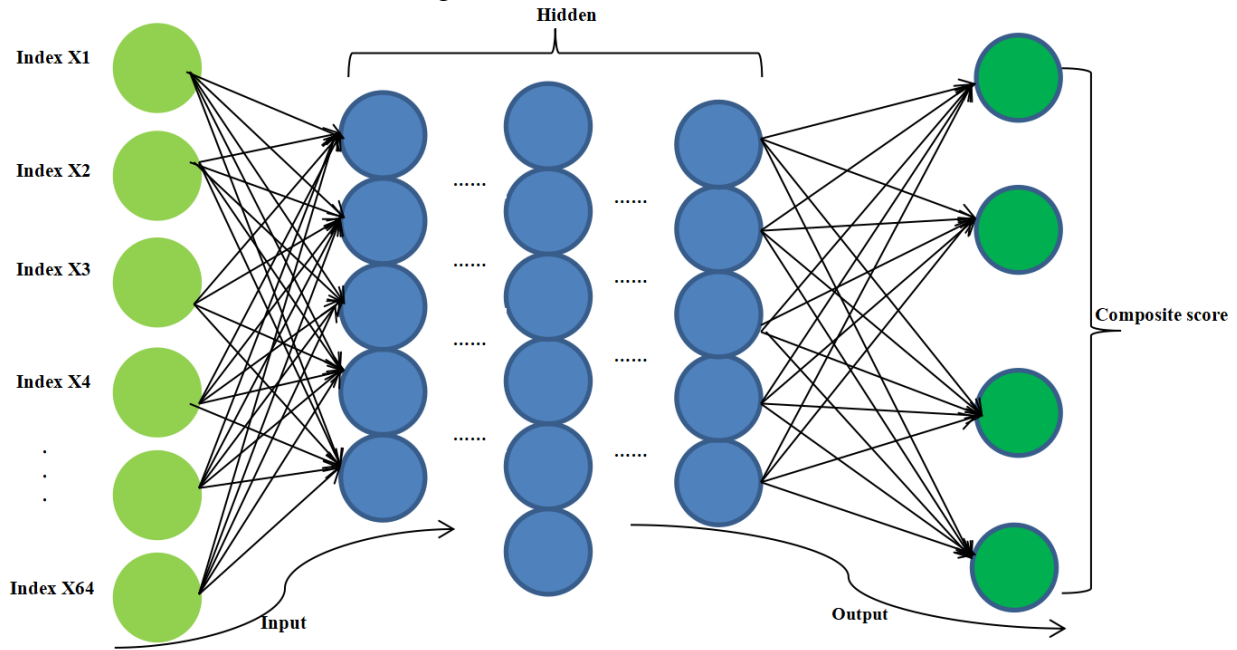

Figure 1 BP neural network structure diagram

The process from the input layer to the hidden layer and from the hidden layer to the output layer is called the forward transfer process. In order to optimize the threshold value of the neural network, the network needs to be adjusted by feedback in order to minimize the error of the output layer[5].

2.2 BPNN model description

After preprocessing the data, the 64 index data were taken as the input set, and the comprehensive score obtained by principal component analysis was taken as the output set to train the neural network. The prediction effect and accuracy were determined by the relative deviation between the actual value and the predicted value and the mean square error.

In this model, a multilayer perceptron function with impulse gradient descent and adaptive learning rate of back propagation is used, that is, traingDX built-in function of MATLAB. The algorithm idea is as follows:

Given the training set,,. $D = \{((x_1, y_1), (x_2, y_2), \cdots, (x_n, y_n))\} x_n \in R^d y_n \in R^l$

Step1: determine the loss function

Suppose the output of BPNN to the sample is (x_k, y_k)

$$\hat{y}_k = (\hat{y}_1^k, \hat{y}_2^k \cdots, \hat{y}_l^k) \tag{2}$$

Among them,. $\hat{y}_j^k = f(\beta_j - \theta_j)$

Then the mean square error of BPNN to the sample is (x_k, y_k)

$$E_k = \frac{1}{2} \sum_{j=1}^{l} (\hat{y}_j^k - y_j^k)^2 \tag{3}$$

Step2: adjustable parameters

Based on the gradient descent method, BPNN algorithm can adjust the parameters of the negative gradient direction of the target.

$$v = v + \Delta v \qquad (4)$$

$$\Delta v = -\eta \frac{\partial E_k}{\partial v} \qquad (5)$$

Step3: calculate the threshold gradient of the output layer
$\hat{y}_j^k$Direct effect, direct effect, using the chain rule$E_k \theta_j \hat{y}_j^k$

$$\frac{\partial E_k}{\partial \theta_j} = \frac{\partial E_k}{\partial \hat{y}_j^k} \frac{\partial \hat{y}_j^k}{\partial \theta_j} \qquad (6)$$

known

$$\frac{\partial E_k}{\partial \hat{y}_j^k} = \hat{y}_j^k - y_j^k \qquad (7)$$

The activation function is the sigmoid function, and can be obtained from this

$$\frac{\partial \hat{y}_j^k}{\partial \theta_j} = -\hat{y}_j^k(1 - \hat{y}_j^k) \qquad (8)$$

Therefore,

$$\frac{\partial E_k}{\partial \theta_j} = -\hat{y}_j^k(1 - \hat{y}_j^k)(\hat{y}_j^k - y_j^k) \qquad (9)$$

Remember to it

$$g_j = \frac{\partial E_k}{\partial \theta_j} \qquad (10)$$

Step4: calculate the gradient of connection weight from the hidden layer to the output layerw_{ij}
$\hat{y}_j^k$Direct effect, direct effect, direct effect$E_k \beta_j \hat{y}_j^k w_{ij} \beta_j$

$$\frac{\partial \beta_j}{\partial w_{kj}} = b_h \qquad (11)$$

$$\frac{\partial E_k}{\partial w_{kj}} = -g_j b_h \qquad (12)$$

Step5: calculate the gradient of hidden layer threshold$\gamma_h \frac{\partial E_k}{\partial \gamma_h}$
Same as above, influence, influence, using the chain rule$b_h E_k \gamma_h b_h$

$$\frac{\partial E_k}{\partial \gamma_h} = b_h(1 - b_h) \sum_{j=1}^{i} g_j w_{hj} \qquad (13)$$

Step6: deduce the results

$$\frac{\partial E_k}{\partial \gamma_h} = b_h(1 - b_h) \sum_{j=1}^{i} g_j w_{hj} \qquad (14)$$

The results show that the threshold gradient of the hidden layer depends on the output of the hidden layer neurons, the threshold gradient of the output layer and the connection weights between them.

2.3 Model Steps

Step1: determine the training set and test set
Training set: 64 index variable data are input set, and the comprehensive score obtained in task 2

is output set.

Prediction set: 64 index variable data is the input set, and the comprehensive score of 5 principal components is the output set.

Step2: train and predict

Firstly, the number of training layers and training times are set, and the neural network training of training set and neural network prediction of prediction set are carried out respectively.

Second, the actual and predicted values of the composite score are output to the same table.

Finally, the number of training layers and training times are changed and repeated until the relative deviation is minimal.

Step3: Find the relative deviation

Compare the predicted values outputed by the test set with those predicted by the training set, and work out the relative deviation between them. The calculation formula of the relative deviation is

$$Pc = \frac{|test-train|}{train} \tag{15}$$

Where, test is the predicted value output by the test set.Train is the predicted data of the training set;Pc is the relative deviation between the two.

Step4: Determine the optimal BPNN prediction model

With relative error as the index, the number of training layers and training times of BPNN algorithm are adjusted to find the number of training layers and training times corresponding to the value of the minimum relative error, which is determined as the optimal BPNN prediction model.

Step5: Confirm the prediction results

The test set is put into the optimal BPNN prediction model for solving.

3. Flow chart of BPNN prediction model

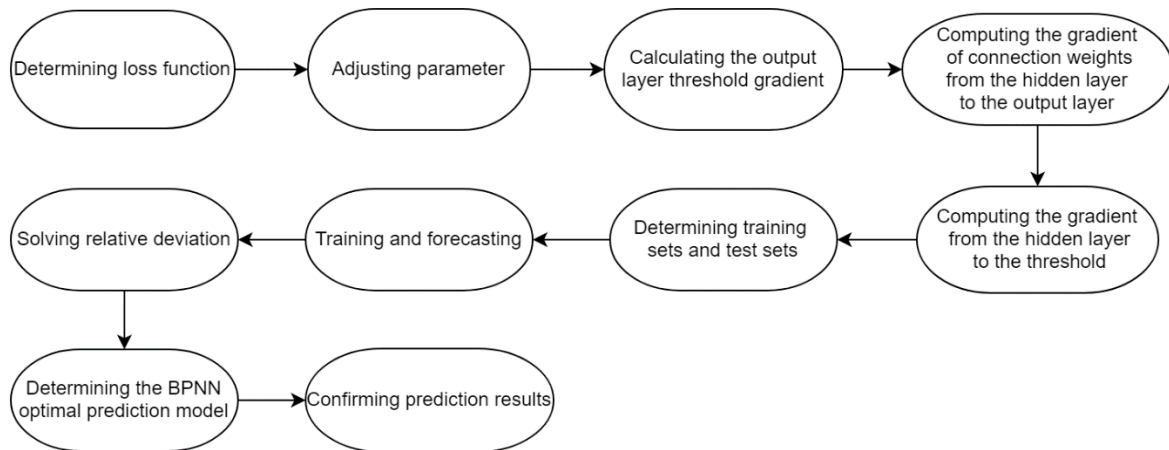The flow chart of BPNN prediction model is shown in Figure 2.

Figure 2 Flow chart of BPNN prediction model

4. Solution of BP neural network model

4.1 Related results of training 10000

The number of training was set to 10000, and the training set and prediction set were substituted into the neural network prediction model. The results obtained by using MATLAB software are shown
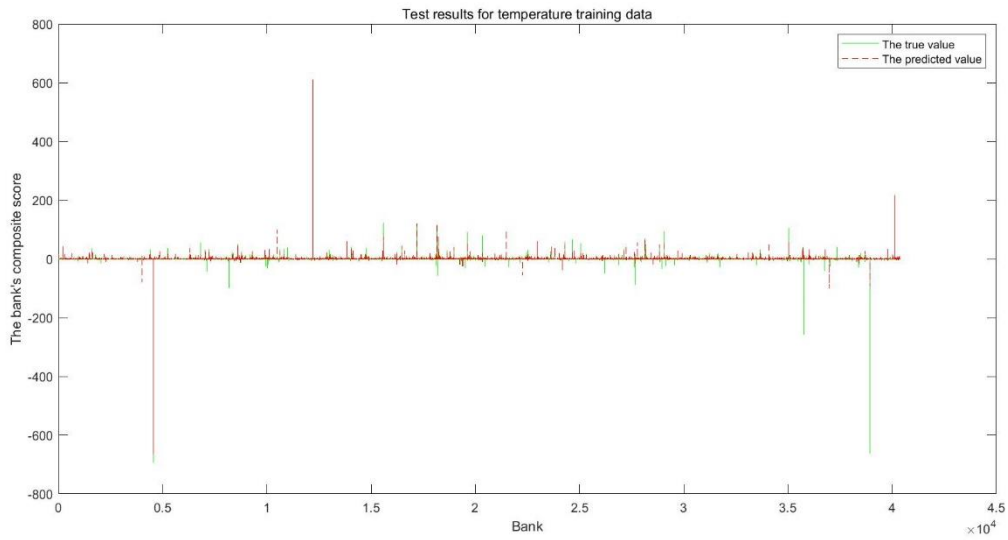
in Figure 3.

Figure 3 Training results

Due to the large data set, the visualization effect needs to be partially magnified and observed. The image processing tool of MATLAB is used to partially enlarge the effect, as shown in Figure 3-4.Through image observation, it can be found that when the number of training reaches 10000, the prediction effect of the model has been better.
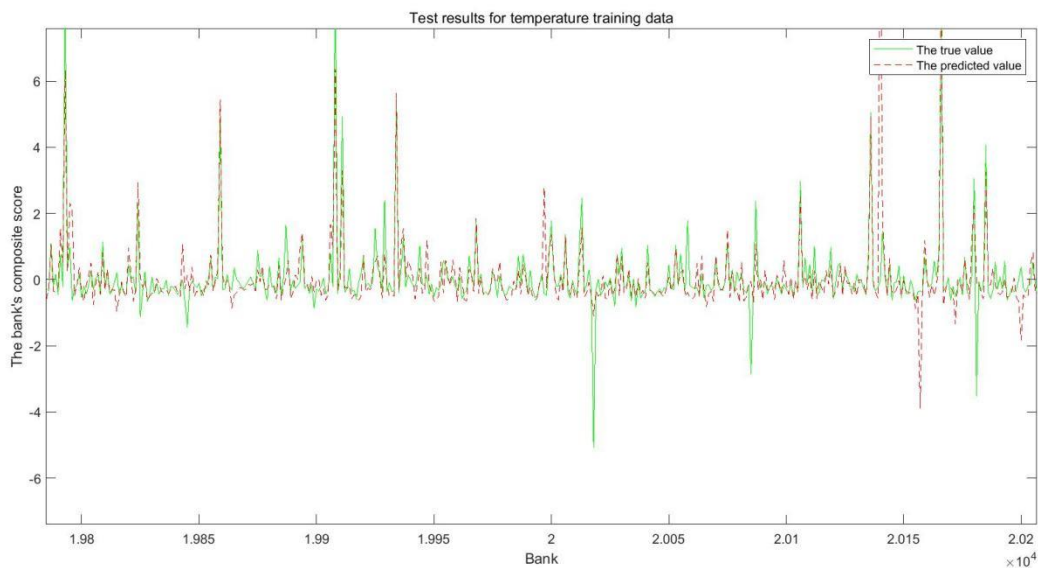

Figure 4. Partial enlargement of visualization effect

4.2 Results of 50,000 training sessions

When the number of training is 50000, the visualization of the real value and predicted value of the comprehensive score is shown in Figure 5.
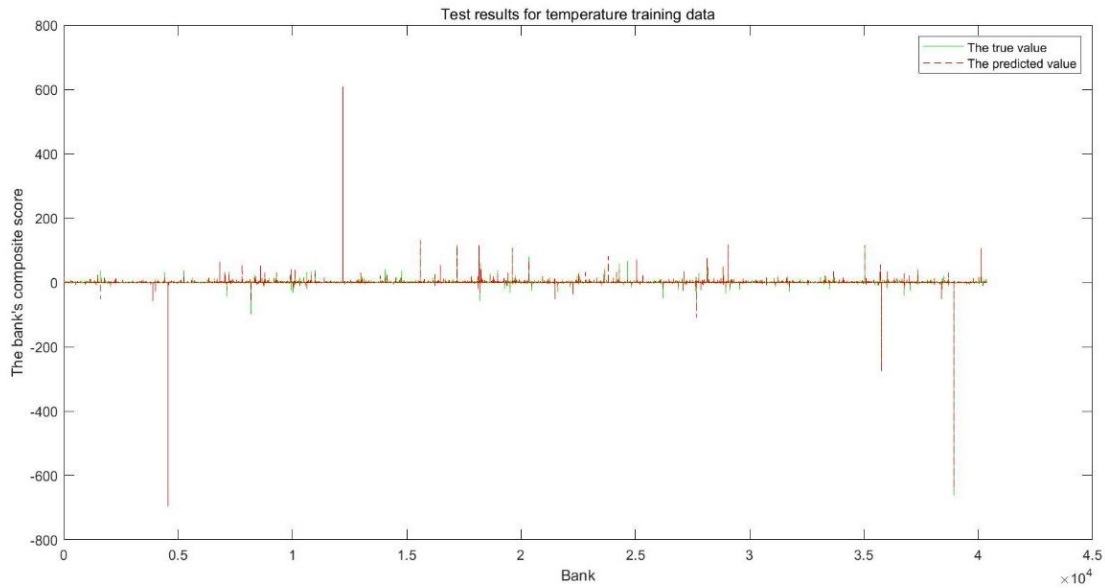
Figure 5 Visualization of Elmannet feedback training at 50000 times

As the data set is too large, the visualization effect needs to be partially magnified and observed. The image processing tool of MATLAB is used, and the partial effect of partial amplification is shown in Figure 6.
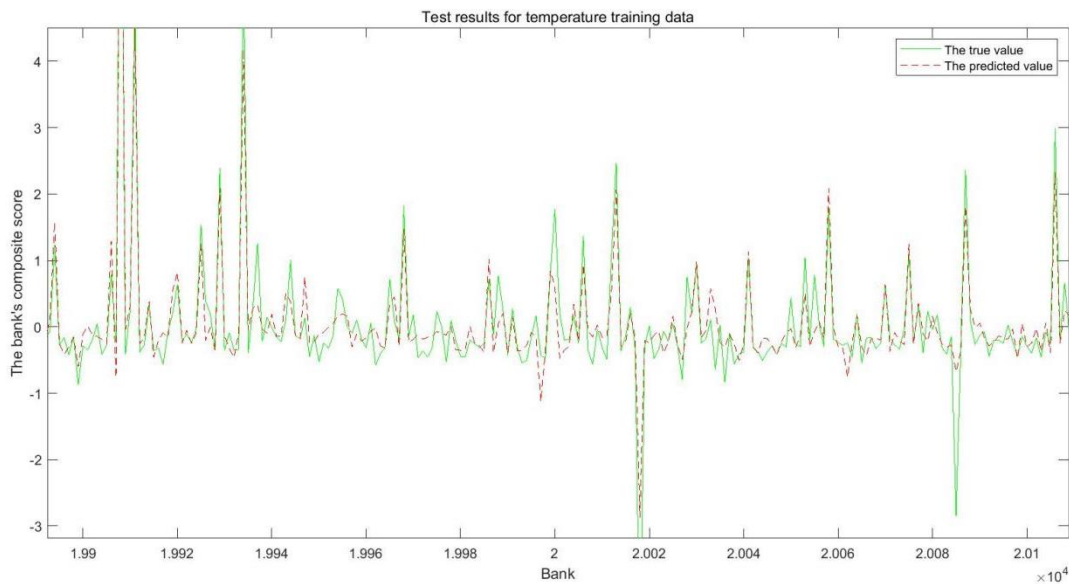

Figure 6 Local shrinkage at 50000 times of feedback training

4.3 Parameter Adjustment Result

When the number of training of BP neural network prediction model is 10000, its mean square error is 14.09038641; when the number of training is 50000, its mean square error is 1.750799366; when the number of training is 100,000, its mean square error is 0.577523428 According to the above adjustment process, it can be found that as the number of training increases, Model accuracy is also rising.Considering the time and model accuracy, the neural network prediction model with 50000

training times is selected in this case. The parameter tuning results are shown in Table 1.

Table 1 Parameter tuning results

Number of training	MSE	Relative error mean
10000	14.09038641	5.193421098
50000	1.750799366	3.656728939
100000	0.577523428	3.675492401

The relative deviation of the three training times is shown in Table 2. Considering the large amount of data, only 20 data are listed in this paper for demonstration.

Table 2 Relative deviation

True value	predict/10000	predict/50000	predict/100,000	Relative Deviation/10000	Relative Deviation/50000	Relative Deviation/100,000
-0.27	2.96	2.75	-0.84	11.91	11.12	2.09
-0.39	3.23	-0.09	-0.74	9.18	0.78	0.88
-0.18	2.76	0.33	-0.56	16.40	2.83	2.11
-0.32	3.06	-0.01	-0.60	10.65	0.97	0.91
-0.30	3.03	0.11	-0.28	11.02	1.37	0.07
-0.36	3.15	0.03	-0.26	9.75	1.09	0.27
-0.17	2.75	0.24	-1.76	16.87	2.40	9.18
13.72	5.02	-0.12	-0.07	0.63	1.01	1.01
-0.31	3.05	0.20	-1.30	10.80	1.63	3.19
-0.12	2.62	1.47	-0.87	23.75	13.71	6.57
0.36	1.62	-1.63	0.39	3.54	5.57	0.09
-0.62	3.74	-0.94	0.47	6.99	0.51	1.75
-0.37	3.19	-0.04	-0.86	9.51	0.88	1.30
-0.63	3.75	-0.72	-0.17	6.96	0.15	0.72
0.09	2.19	0.55	-2.33	23.71	5.23	27.30
0.34	1.65	-1.51	0.63	3.85	5.42	0.83
-0.36	3.15	-0.20	-0.69	9.78	0.43	0.93
-0.30	3.03	-0.85	-0.34	11.04	1.82	0.12
-0.21	2.83	-0.71	-0.17	14.32	2.32	0.21
-0.21	2.82	-0.04	-1.13	14.65	0.83	4.46

5. Model solution results

Through the analysis of the diagram of the model solving process, the following conclusions can be drawn:

(1) Under different training times, each sample data comprehensive scoring the real value of data and predicted data reflects the fitting effect of the established BP neural network prediction model of prediction effect, at the same time, from the visual diagram can be found with the increase of the number of training, the accuracy of the prediction is also on the rise, but also increase the time needed for training.

(2) The relative deviation between the predicted value and the actual value of the comprehensive score of each bank is small, so it is found that the model error is small.

(3) Through the established BP neural network prediction model, for the existing Banks in Poland, the future development of their banks has a fairly good prediction effect. Thus make business plan and personnel management adjustment. The visualization between predicted and actual values is shown in Figure 7.
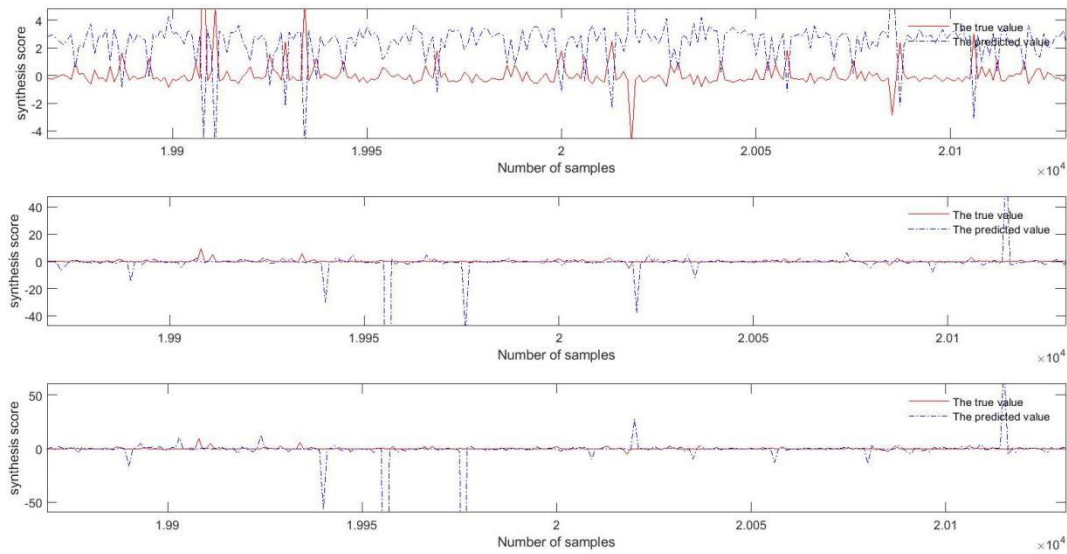
Figure 7 Visualization of predicted and actual values

References

[1] Liu Lulu, Ning Xin. A Self-Learning BP Neural Network Assessment Algorithm for Credit Risk of Commercial Bank[J]. Wireless Communications and Mobile Computing, 222, 2022.

[2] Wang Xiaogang. Analysis of Bank Credit Risk Evaluation Model Based on BP Neural Network.[J]. Computational Intelligence and neuroscience, 2022,202, 2.

[3] Shan Zhirui. Credit Risk Assessment and Prediction Model of Commercial Banks[J]. Financial Engineering and Risk Management, 2021, 4 (4).

[4] Fariyanti. Iskandar, Malani Rheo, Suprapty Bedi. Total asset prediction of the large Indonesian bank using adaptive artificial neural network back-propagation[J]. International Journal of Engineering & Technology, 2018, 7 (2.2).

[5] Honglei Zhang, Wenshan Yuan, Hua Jiang. Performance Evaluation On Human Resource Management Of China'S Commercial Banks on Improved Bp Neural Networks [J]. International Journal of Advancements in Computing Technology, 2012, 4 (11).