Portfolio Research Based on Neural Network and Multi-Objective Programming

Rongjie Jian
Xiamen Huaxia University
Xiamen, China 361024

Tingting Ni
Xiamen Xingcai Vocational&Technical College
Xiamen, China 361021

Abstract—**In financial investment, risk and benefit coexist. How to balance the benefits and risks and find the optimal investment portfolio is a key issue to be considered by investors. In this research, BP neural network is used to predict the future return on equity (ROE) of asset; a multi-objective programming model of investment portfolio is established on the basis of Markowitz's portfolio investment theory; to select the optimal investment portfolio, a comparison is made on the benefit-risk ratio of the investment portfolio with the smallest risk and at different income levels; in addition, an empirical analysis is made on the basis of the quarterly ROE data of 5 stocks during 2002-2017.**

Keywords—*portfolio; BP neural network; ROE (return on equity); multi-objective programming; benefit–risk ratio*

I. Introduction

A portfolio is a collection of stocks, bonds, financial derivative products and the like held by investors. When making investment, rational investor often may not put "eggs" in the same basket. In order to spread the risk, they may make a portfolio investment [1]. In financial investment, investors always want to find a portfolio with the largest profit and the least risk, namely an optimal portfolio. However, benefit and risk always follow up like body and its shadow; they are generally positively correlated. The maximum benefit and minimum risk are often contradictory [2]. Therefore, weighing the benefit and risk to find the optimal investment portfolio is a key problem to be considered in making investment decision. With respect to this problem, there have already been a plenty of researches, and different multi-objective programming models have been established from different aspects to solve the problem [3] [4] [5] [6] [7] [8]; but most of the models are based on the historical data of the rate of return to assess the benefit and risks; few of them use predicted rate of return to make the assessment; and further weighing of benefit and risk is even not made. There are many common data prediction methods, such as moving average method, exponential smoothing method, regression prediction method, gray prediction method and neural network [9] [10] [11] [12] [13]. The rate of return is generally nonlinear and fluctuating. Because neural network has strong nonlinear adaptability [14], in this research, the most widely used BP neural network is used to predict the rate of return; a multi-objective programming model for solution is established on the basis

of Markowitz's portfolio investment theory [15]; a calculation and comparison is made on the benefit-risk ratio of the investment portfolio with the smallest risk and at different income levels; and finally, the optimal investment portfolio is found.

II. BP Neural Network

A. Overview of BP Neural Network

The BP neural network is a multi-layer feed-forward network composed of input layer, output layer and at least one hidden layer. Its learning process can be divided into two phases: information forward transfer and error back propagation. In the information forward transfer phase, information is first input into the network by the input layer, then processed and transferred layer by layer, and finally reaches the output layer. If the output layer does not reach the ideal output at this time, the information will enter the error back propagation stage, and the weight and threshold of the network will then be adjusted according to the error of the output layer and the ideal output. Such adjustment should be made over and over again to make the output layer closer and closer to the ideal output till the error between them meets certain accuracy requirements [16]. The learning process of BP neural network is as shown in "Fig. 1".
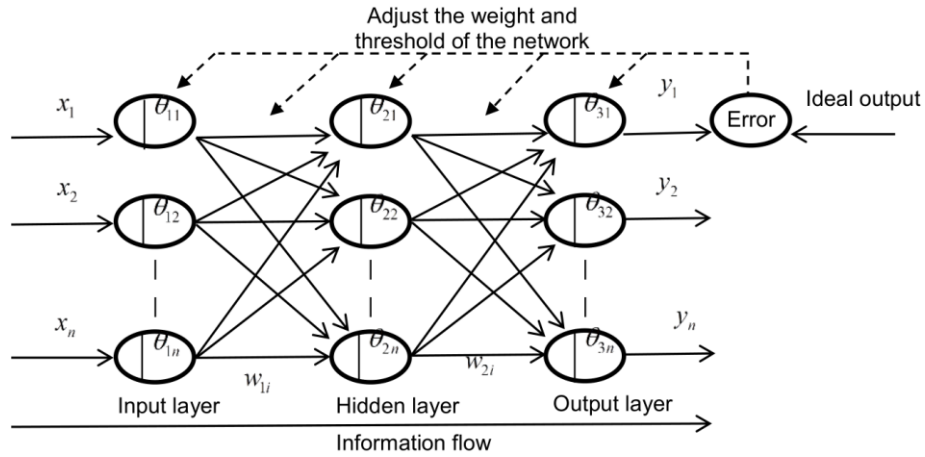
Fig. 1. Learning process of BP neural network.

Wherein, $x_i(i=1,2,\ldots,n)$ is input information; $y_i(i=1,2,\ldots,n)$ is output information; $w_{1i}, w_{2i}(i=1,2,\ldots,n)$ is the link weight between layers; $\theta_{1i}, \theta_{2i}, \theta_{3i}(i=1,2,\ldots,n)$ is the threshold in the network.

B. Creation of BP Neural Network

1) Data preprocessing:

Before creating a BP neural network, it is needed to standardize the data and transform the data into the interval [0, 1] in order to speed up the convergence of the neural network. The transformation formula is as follows:

$$\hat{x}_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \qquad (1)$$

Among them, $x_{\max}$ and $x_{\min}$ respectively represent the maximum and minimum values among the data.

2) Structure of BP neural network:

As the 3-layer neural network can approximate any complex nonlinear function with arbitrary precision, a 3-layer BP neural network with relatively simple structure is selected in this research [17].

3) Determination of the number of nodes at input/output layer:

The number of nodes at input/output layer is determined by the number of input/output data. For example, if n data are input, the number of nodes at input layer is n; if m data are output, the number of nodes at output layer is m.

4) Determination of the number of nodes at hidden layer:

According to Kolmogorov's theorem [18], for a 3-layer BP neural network, the number of nodes at hidden layer is 2n+1; wherein, n is the number of nodes at input layer.

5) Selection of activation function:

There are two purposes for setting the neural network activation function: one is to enable the neural network to approximate the nonlinear function with arbitrary precision; the other is to map the output value of the neuron to the bounded cell and speed up the convergence of the neural network.

BP neural network has many kinds of activation functions. Commonly, there are piecewise linear functions, log-sigmoid functions and tan-sigmoid functions. The input values of the three common activation functions can be any values. The main difference is the mapped interval of output value and the smoothness of the function graph. The piecewise linear function maps the output value to the interval [0,1] and the function graph is not smooth; the log-sigmoid function maps the output value to the interval (0,1) and the function graph is smooth; the tan-sigmoid function maps the output value to the interval (-1, 1) and the function graph is smooth. Graphs of common activation functions are as shown in "Fig. 2".
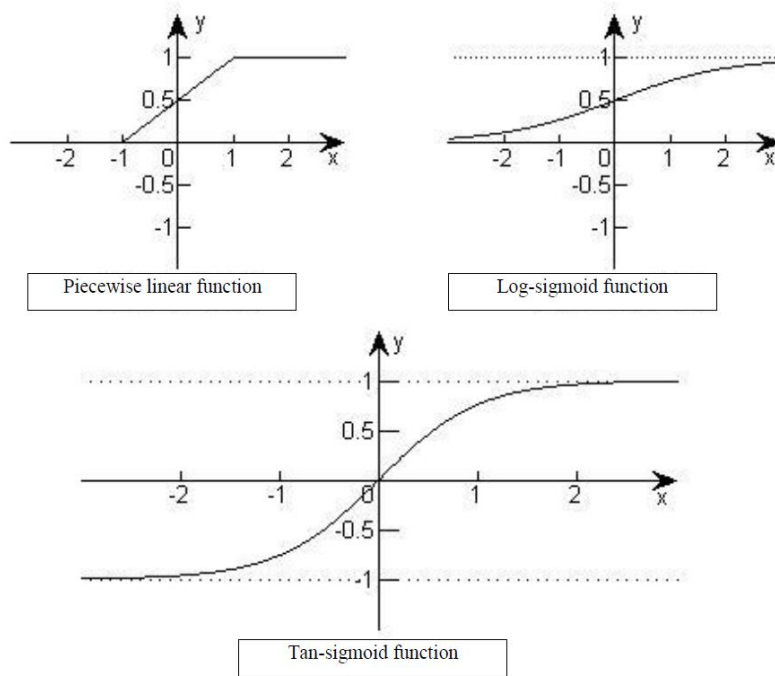
Fig. 2. Graphs of common activation functions.

III. MULTI-OBJECTIVE PROGRAMMING MODEL

A. Building of Multi-objective Programming Model

Multi-objective programming is a branch of mathematical programming and aims at finding the optimal solution in the given region when there is more than one objective function. When selecting a portfolio, it is always expected to get the most benefit at the least risk; hence, the selection of portfolio is a multi-objective programming problem.

It is assumed that there are n kinds of assets $s_i (i=1,2,\ldots,n)$ available for investors to choose from; the rate of return of asset s_i is r_i, the standard deviation is σ_i, and the proportion of the investment asset s_i is w_i. At this point, rational investors may make portfolio investment, both expecting to obtain the largest benefit and bear the least risk. The largest benefit and the least risk are two goals. Therefore, a multi-objective programming model can be established on the basis of Markowitz's portfolio investment theory.

1) Objective function:
Objective function 1: The largest benefit, namely

$$\max \sum_{i=1}^{n} r_i w_i$$

Objective function 2: The least risk, namely

$$\min \sum_{i=1}^{n}\sum_{i=1}^{n} \text{cov}(r_i, r_j) w_i w_j$$

Among them, $\text{cov}(r_i, r_j)$ is the covariance of the returns of assets s_i and s_j.

2) Constraints:
Constraint 1: Only invest in the *n* kinds of assets, namely

$$\sum_{i=1}^{n} w_i = 1$$

Constraint 2: Each asset is not allowed to be in short selling status, namely $w_i \geq 0 (i=1,2,\ldots,n)$

3) Multi-objective programming model:
Combined with objective functions 1 and 2 and constraints 1 and 2, it is possible to build a multi-objective programming model of investment portfolio:

$$\begin{cases} \max \sum_{i=1}^{n} r_i w_i \\ \min \sum_{i=1}^{n}\sum_{i=1}^{n} \text{cov}(r_i, r_j) w_i w_j \\ \sum_{i=1}^{n} w_i = 1 \\ w_i \geq 0, i=1,2,\ldots,n \end{cases} \quad (2)$$

B. Simplification of Multi-objective Programming Model: Quadratic Programming Model

Obviously, the multi-objective programming model established has two objective functions; but these two objectives are often opposite. To facilitate the solution, the model is simplified, namely the benefit level is fixed and the

risk is minimized. If the investor wants the return level to be K, the portfolio with the least risk can be found at this time and the corresponding nonlinear programming model is:

$$
\begin{cases}
\min \sum_{i=1}^{n} \sum_{i=1}^{n} \text{cov}\left(r_i, r_j\right) w_i w_j \\
\sum_{i=1}^{n} w_i = 1 \\
w_i \geq 0, i = 1, 2, \ldots, n \\
\sum_{i=1}^{n} r_i w_i = K
\end{cases}
\tag{3}
$$

If the objective function of the nonlinear programming is the quadratic function of independent variable w and the constraints are all linear, this nonlinear programming is called quadratic programming. As can be seen from equation (3), the simplified model is a quadratic programming model.

IV. BENEFIT-RISK RATIO

When the benefit level K is changed, the portfolio with the least risk at different benefit levels can be acquired after solving the said quadratic programming model. At this point, it is needed to further measure the benefit and risk to find the optimal portfolio; therefore, it is needed to introduce the benefit-risk ratio.

Benefit-risk ratio is the ratio of possible benefits to possible losses and measures the benefit brought about by unit loss. If the larger the ratio is, the larger the benefit will be brought about by unit loss, it indicates that the portfolio is good; otherwise, the portfolio is poor; and the portfolio corresponding to the maximum value of the ratio is the optimal portfolio. In order to make better use of the benefit-risk ratio, it is feasible to replace possible benefit with benefit level and replace possible loss with risk; then the benefit-risk ratio is:

$$
\frac{\sum_{i=1}^{n} r_i w_i}{\sum_{i=1}^{n} \sum_{i=1}^{n} \text{cov}\left(r_i, r_j\right) w_i w_j}
\tag{4}
$$

Among them, $\sum_{i=1}^{n} r_i w_i$ is the benefit level of the portfolio, and $\sum_{i=1}^{n} \sum_{i=1}^{n} \text{cov}\left(r_i, r_j\right) w_i w_j$ is the risk of the portfolio.

V. EMPIRICAL ANALYSIS

The data in this paper is sourced from a Software Flush; 5 stocks with good investment prospects are selected from the A-share market as investable assets, namely Bank of China (601988), Hengtong Optic-electric (600487), Guangdong Expressway A (000429), Yangtze Power (600900), and Zhifei Bio (300122). The quarterly ROE of the 5 stocks from 2002 to 2017 are selected as the analysis data.

A. Assessment of Benefits and Risks

If the previous ten periods of quarterly ROE are used to predict the future five periods of quarterly ROE, the numbers of neurons at the input and output layers are respectively 10 and 5. The number of neurons at the hidden layer can be determined as 21 according to Kolmogorov's theorem. A relatively simple 3-layer neural network structure can be adopted; the activation functions of the input layer and the output layer can select the log-sigmoid function; and the activation function of the hidden layer can select the tan-sigmoid function.

Matlab2016a software can be used to separately train five BP neural networks. Taking quarterly ROEs of each stock in the last ten periods as input layer data, the corresponding BP neural network can be used to predict the quarterly ROE in the future five periods. The latest ten periods of quarterly ROEs of each stock and the predicted values in the future five periods are respectively shown in "Table I" and "Table II".

TABLE I. THE LATEST TEN PERIODS OF QUARTERLY ROEs OF EACH STOCK (UNIT: %)

Stock name	Recent ten periods of quarterly ROE									
Bank of China	2.68	3.93	3.21	1.96	2.80	3.44	3.46	2.64	3.01	3.70
Hengtong Optic-electric	4.16	7.79	4.50	4.02	10.70	5.38	2.40	3.29	5.73	2.18
Guangdong Expressway A	3.91	7.30	3.33	2.18	3.38	4.09	2.44	1.22	3.36	2.65
Yangtze Power	6.91	4.65	2.06	3.03	6.47	4.68	2.03	1.89	4.64	3.94
Zhifei Bio	3.80	4.08	2.34	1.02	-0.21	-1.86	2.34	1.42	2.23	2.44

TABLE II. PREDICTED QUARTERLY ROEs OF EACH STOCK IN THE FUTURE FIVE PERIODS (UNIT: %)

Stock name	Predicted value				
Bank of China	4.4759	4.4838	6.8953	4.5497	6.1977
Hengtong Optic-electric	2.7746	2.7087	1.9214	1.7442	1.5974
Guangdong Expressway A	1.7849	1.7871	1.8803	2.3146	1.7675
Yangtze Power	3.0949	3.0008	4.5055	5.6315	3.0246
Zhifei Bio	1.5700	1.5700	1.5831	3.1524	3.1046

Taking the average of the predicted quarterly ROEs of each stock in the future five periods as the benefit, the covariance between them as the risk, it is possible to evaluate and get the benefit and risk of each stock, as shown in "Table III" and "Table IV".

TABLE III.   BENEFIT AND VARIANCE OF EACH STOCK (UNIT: %)

Stock name	Benefit	Variance
Bank of China	5.3205	1.3143
Hengtong Optic-electric	2.1493	0.3061
Guangdong Expressway A	1.9069	0.0539
Yangtze Power	3.8515	1.3940
Zhifei Bio	2.1960	0.7249

TABLE IV.   COVARIANCE BETWEEN THE STOCKS (UNIT: %)

Stock portfolio	(s_1, s_2)	(s_1, s_3)	(s_1, s_4)	(s_1, s_5)	(s_2, s_3)
Covariance	-0.3817	-0.0688	0.0708	0.0368	-0.0564
Stock portfolio	(s_2, s_4)	(s_2, s_5)	(s_3, s_4)	(s_3, s_5)	(s_4, s_5)
Covariance	-0.3407	-0.3727	0.2544	0.1077	0.3891

In "Table IV", s_1 represents stock of the Bank of China; s_2 represents stock of the Hengtong Optic-electric; s_3 represents stock of Guangdong Expressway A; s_4 represents stock of Yangtze Power; s_5 represents stock of Zhifei Bio.

As can be seen from "Table III", in terms of a single stock, the stock of the Bank of China has the highest benefit, reaching 5.3205%, and the stock of Guangdong Expressway A has the lowest benefit of only 1.9069%; the stock of Changjiang Power has the highest variance, reaching 1.3940%. The stock of Guangdong Expressway A has the least risk, i.e., 0.0539%; although the stock of Yangtze Power has relatively high benefit, its variance is the largest and its benefit is very unstable; the stock of Guangdong Expressway A has the least benefit, but its variance is the least.

As can be seen from "Table IV", the covariance of stock portfolios of Bank of China and Hengtong Optic-electric,

Bank of China and Guangdong Expressway A, Hengtong Optic-electric and Guangdong Expressway A, Hengtong Optic-electric and Yangtze Power, Hengtong Optic-electric and Zhifei Bio are negative; this indicates that investing in those five stock portfolios can reduce investment risk. The covariance of the remained stock portfolios is positive; this indicates that investing in the remained stock portfolios may increase the investment risk. In particular, the covariance of the stock portfolio of Bank of China and Hengtong Optic-electric is the lowest negative value, reaching -0.3817%, and the covariance of Yangtze Power and Zhifei Bio is the largest positive value, reaching 0.3891%; this indicates that, compared with other stock portfolios, the stock portfolio of Bank of China and Hengtong Optic-electric can reduce the risk the most, while the stock portfolio of Yangtze Power and Zhifei Bio may increase the risk the most.

B. Building and Solution of Quadratic Programming Model of Investment Portfolio

The assessed benefit and risk is used to fix the benefit level K and minimize the risk; on this basis, it is possible to build the quadratic programming model of the investment portfolio:

$$
\begin{cases}
\min \sum_{i=1}^{5} \sum_{i=1}^{5} \text{cov}(r_i, r_j) w_i w_j \\
\sum_{i=1}^{5} w_i = 1 \\
w_i \geq 0, i = 1, 2, \ldots, 5 \\
\sum_{i=1}^{5} r_i w_i = K
\end{cases} \tag{5}
$$

Wherein, $w_i (i = 1, 2, \ldots, n)$ is the investment proportion of the i th type of stock.

The software Matlab2016a is used to solve the quadratic programming model presented by equation (5), and the minimum risk variation curve at different benefit levels is obtained as shown in "Fig. 3".
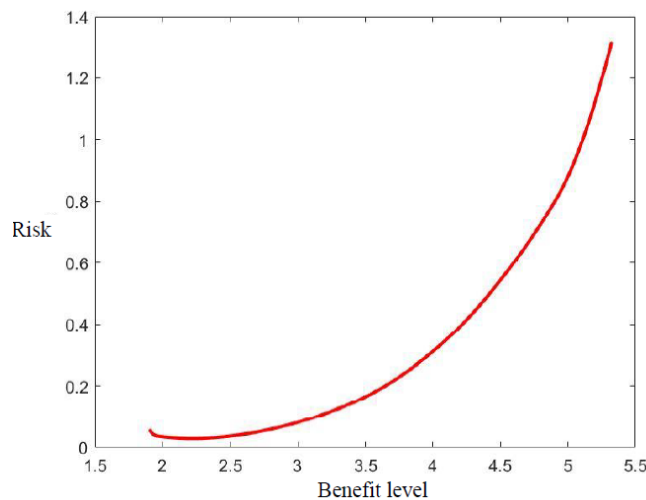

Fig. 3.   Curve minimum risk variation at different benefit levels.

As can be seen from "Fig. 3", the curve first lowers and then rises. When the benefit level reaches 2.2202%, the risk reaches the minimum value, 0.0289%. At this time, the corresponding investment proportion of each stock is as shown in "Table V".

TABLE V.   INVESTMENT PROPORTION OF EACH STOCK

Stock name	Bank of China	Hengtong Optic-electric	Guangdong Expressway A	Yangtze Power	Zhifei Bio
Investment proportion	7.15%	22.82%	65.21%	0%	4.82%

C. Selecting the Optimal Investment Portfolio

Although when the return level is 2.2202% and the risk is 0.0289%, the portfolio's benefit-risk ratio 76.82 is already quite large, the portfolio at this time may not be the optimal one. In order to find the optimal portfolio, it is needed to calculate the benefit-risk ratio of the portfolio with minimum risk at different benefit levels; the portfolio with the largest benefit-risk ratio is the optimal portfolio. The benefit-risk ratio curve of each portfolio is shown in "Fig. 4".
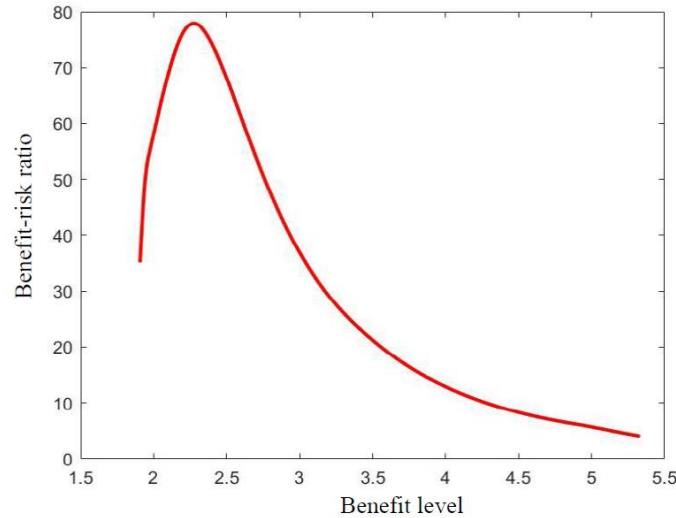

Fig. 4.   Benefit-risk ratio curve of each portfolio.

As can be seen from "Fig. 4", the benefit-risk ratio curve first rises and then lowers. When the rate of return reaches 2.2760%, the benefit-risk ratio reaches the maximum value, 77.9103. At this time, the portfolio is the optimal portfolio, and the investment proportion of the corresponding stock is shown in "Table VI".

TABLE VI.   INVESTMENT PROPORTION IN EACH STOCK IN THE OPTIMAL INVESTMENT PORTFOLIO

Stock name	Bank of China	Hengtong Optic-electric	Guangdong Expressway A	Yangtze Power	Zhifei Bio
Investment proportion	8.70%	23.60%	62.54%	0%	5.16%

VI.   CONCLUSION

Taking stock investment as an example and according to the ROE of each stock in the previous periods, this paper uses BP neural network to predict the quarterly ROE of the future five periods, and builds up a multi-objective programming model taking the average value as the benefit and the covariance as the risk. To facilitate solving it, fixing the benefit level and minimizing the risk, the multi-objective programming model is transformed into a quadratic programming model. By solving the quadratic programming model, the investment portfolio with the lowest risk at different benefit levels is obtained, and by further calculating the benefit-risk ratio of each portfolio, the optimal investment portfolio is finally found.

REFERENCES

[1]   Yan Fei, Zhang Chongqi. Experimental Design Method for Portfolio Investment [J]. Journal of Applied Statistics and Management, 2016, 35(1). (in Chinese)

[2]   Huang Dongbin, Wang Yong, Liu Qiyan. Portfolio Choices with Entropy-TOPSIS Stocks Pre-evaluation [J]. Review of Investment Studies, 2017(2). (in Chinese)

[3]   Tong Wenbing. Building and Solution of Linear Programming Model for Risk Investment Portfolio [J]. Statistics & Decision, 2016(9). (in Chinese)

[4]   Liu Menghui. Optimal Life Insurance Investment Portfolio Model Based on Nonlinear Multi-objective Programming [J]. Statistics & Decision, 2011(19). (in Chinese)

[5]   Hu Dasha, Wu Wei. The Application of Goal Programming Method to the Portfolio Investment [J]. Operations Research and Management Science, 2004, 13(3). (in Chinese)

[6] Wang Mengdong, Tong Shikuan. Multi-objected Investment Combination Model Based on Quadratic Programming [J]. Journal of Wuhan University of Technology, 2007, 29(8). (in Chinese)

[7] Wu Wei, Cao Yanfei. Research on Target Planning Model of Securities Investment Portfolio [J]. Statistics & Decision, 2010(18). (in Chinese)

[8] Zhao Yumei, Chen Huayou. Interval Number Linear Programming Model of Portfolio Investment [J]. Operations Research and Management Science, 2006, 15(2). (in Chinese)

[9] Zou Jiaojiao, Yang Xiaobing, Kong Deguang. Epidemic characteristics and tendency prediction of hand foot mouth disease in Wuhan (2011-2016) [J]. Journal of Public Health and Preventive Medicine, 2017, 28(4). (in Chinese)

[10] Gong Wenwei, Huang Jing. A Demand Forecast Model Based on the Gray Theory and Eexponential Smoothing Method [J]. Statistics & Decision, 2017(1). (in Chinese)

[11] Yang Huaide, Feng Qi, Guo Xiaoyan. Analysis on the variation of groundwater level and its influence factors in Minqin Oasis based on the regression model [J]. Journal of Arid Land Resources and Environment, 2017,31(2). (in Chinese)

[12] Lam M. Neural Network Techniques for Financial Performance Prediction: Integration Fundamental and Technical Analysis [J]. Decision Support Systems, 2004, 37(4).

[13] Shen K. Implementing Value Investing Strategy by Artificial Neural Network [J]. International & Information Technology, 2011, 1 (1).

[14] Li Bo, Wang Jianguo,Li Jingwen. Prosperity Prediction Model Based on Neural Network [J]. Statistics & Decision, 2017(10). (in Chinese)

[15] Markowitz H. Portfolio Selection [J]. Journal of Finance, 1952, 7 (1).

[16] Shi Feng. 30-case Analysis of Matlab Neural Network [M]. Beijing: Beihang University Press, 2010. (in Chinese)

[17] Xue Ye, Yan Qizhu, Ren Yao. Research on Inflation Risk Forecast of China Based on the Big-data Technology [J]. On Economic Problems, 2016(1). (in Chinese)

[18] Liu Ougeng, Li Shengqing, Xiao Qianghui. Research on A Structure of Multi-Layer Forward Artificial Neural Network [J]. Journal of Natural Science of Hunan Normal University, 2004,27(1). (in Chinese)