

SHORT TERM STOCK PRICE FORECASTING WITH APPLICATION OF NEURAL  
NETWORK

by

Jiayin Liu

Bachelor's degree, Ritsumeikan Asia Pacific University, Japan, September 2010

A Project,  
presented to Ryerson University

in partial fulfillment of the  
requirements for the degree of  
Master of Engineering  
in the Program of  
Electrical and Computer Engineering

Toronto, Ontario, Canada, 2015

©(Jiayin LIU) 2015

## AUTHOR'S DECLARATION FOR ELECTRONIC SUBMISSION OF A PROJECT

I hereby declare that I am the sole author of this project. This is a true copy of the project, including any required final revisions.

I authorize Ryerson University to lend this project to other institutions or individuals for the purpose of scholarly research

I further authorize Ryerson University to reproduce this project by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

I understand that my thesis may be made electronically available to the public.

# Short Term Stock Price Forecasting with Application of Neural Network

Master of Engineering2015

Jiayin LIU

Electrical and Computer Engineering

Ryerson University

## Abstract:

With the world's rapid economic growth and the expansion of stock market, It produced a large amount of valuable data information. That data become an important investors in stock investment analysis subject. Thorough analysis the short-term stock price forecast problem and comparing a variety of stock price forecasting method, on the basis of BP neural network(BPNN) <sup>[1]</sup> and principal component analysis (PCA)<sup>[2]</sup> and genetic algorithm and the feasibility of short-term prediction of stock price. BP neural network can use the study of historical stock market data, find out the inherent law of development and change of the stock market, so as to realize the future stock price data changes over a period of time.

Key words: Stock price prediction; The BP neural network; Principal component analysis; Genetic algorithm

## Acknowledgments:

I would like to express my gratitude to my supervisor Dr.Kaamran Raahemifar for the useful comments, remarks and engagement through the learning process of this master thesis. Furthermore, I would like to thank the participants in my survey, who have willingly shared their precious time and the support on the way.Also,I would like to thank my parents and friends, who have supported me throughout entire process, both by keeping me harmonious and helping me putting pieces together.

## Table of Contents:

Title Page	
Author's declaration for electronic submission of a MRP	ii
Abstract	iii
Acknowledgments	iv
Table of Contents	v
List of tables	viii
List of figures	ix
List of symbol	x
List of appendices	xi
Chapter 1 The main research contents and thesis structure	01
Chapter 2 Price short-term prediction fundamental theory and method	03
2.1 Factors of affecting stock prices	03
2.2 The difficulty to predict the price of stock	05
2.3 Stock price forecasting technology methods	05
2.3.1 The traditional analysis method	05
2.3.2 Time series analysis method	07
2.3.3 NN prediction method	08
2.3.4 Several forecasting methods	09
2.4 Summery of this chapter	10
Chapter 3 Principal component analysis and BPNN	11
3.1 The characteristics of the BPNN	14
3.2 BP NN model in stock price advantages &problem in short-term prediction	16
3.3 Dimension about the BPNN	17
3.4 Dimension about reduction of the commonly used method	17
3.5 Introduction of PCA	18
3.6 Relevant principle of PCA	19

3.7 Summary of this chapter	22
Chapter 4 Optimize BPNN model based on the genetic algorithm	23
4.1 Concept and principle of genetic algorithm	23
4.1.1 Basic concept of GA	23
4.1.2 Principle of GA and basic operation	26
4.1.3 Process of GA	33
4.1.4 The advantages of GA	36
4.2 Genetic neural network model	37
4.2.1 The genetic neural network concept	37
4.2.2 The combination of GA and NN	38
4.2.2.1 GA optimization of connection weight of NN	38
4.2.2.2 GA the optimization of NN structure	40
4.2.2.3 GA to optimize the NN learning rules	40
4.2.3 Construction of a genetic neural network model	42
4.3 Summary of this chapter	45
Chapter 5 Realization and experiment of composite model of stock price forecast	46
5.1 Model design	46
5.1.1 Overall design model	46
5.1.2 Summary of this part	47
5.2 Experiment and result analysis	48
5.2.1 Performance evaluation index	48
5.2.2 Dimension reduction methods about contrast experiment	48
5.2.2.1 Experiment index selection	49
5.2.2.2 Principal component analysis of optimization for NN model	51
5.2.2.3 Related principal component analysis to optimize BP NN prediction model	54
5.2.2.4 Error analysis	56
5.3 Genetic algorithm to optimize the BP neural network contrast experiment	56

5.3.1 Genetic algorithm to optimize the BP neural network training process	57
5.3.2 BP neural network training process	58
5.3.3 Error analysis	62
5.4 Comprehensive prediction model experiment	63
5.4.1 Results	63
5.4.2 Error analysis	64
5.5 Summary of this chapter	65
Chapter 6 Summary and outlook	66
6.1 Summary	66
6.2 Shortness and prospect	67
Appendices	69
Reference	81

## List of Tables

Tab 5.1 The Experiment selected Variable.	51
Tab 5.2 Select the training results of comparison number of neurons in the hidden layer	53
Tab 5.3 Comparison of PCA of BPNN forecasting results and actual value	54
Tab 5.4 Select training results of comparison number of neurons in the hidden layer	55
Tab5.5 Related principal component analysis predicted results comparison of BP NN	55
Tab.5.6 Comparison of the prediction error of the different number of initial population table	57
Tab.5.7 Kinds of improved BP algorithm forecast error comparison	62
Tab 5.8 ~Tab 5.10 Analysis the relate PCA;PCA;GA_BP algorithm result	63~64
Tab 5.11 Comparison table of forecast model error	65

## List of Figures

Fig 3.0 Graph of corresponding function	12
Fig3.1 Three layer BPNN	14
Fig 3.2 The basic structural diagram of the BP prediction model	16
Fig 3.3 Venn diagram of good of fit	20
Fig 4.0 New individual combination	25
Fig 4.1 Gene code chain variation	25
Fig 4.2 Two points crossover	30
Fig 4.3 Uniform crossover	31
Fig 4.4 Genetic algorithm flowchart	34
Fig 4.5 Genetic algorithm flowchart	35
Fig 4.6 Neural Network model flowchart	44
Fig 5.1 Comprehensive prediction model	47
Fig 5.2 The experimental results comparison chart	56
Fig 5.3 ~ Fig 5.9 Kinds of improved BPNN function	60~62
Fig 5.10 The experimental result comparison chart	64

## List of symbol

PCA	Principal Component Analysis
RBF	Radial Basis Function
BPNN	Error Back Propagation Neural Network
SOM	Self-organizing Map
GA	Genetic Algorithm
NN	Neural Network
ANN	Artificial Neural Network
SE	Squares' Sum Error
LM	Levenberg Marquardt
MSE	Mean Square Error
MA	Moving Average
WR	William Indicators
K,D,J	Stochastic
BIAS	Bias rate
PSY	Physiological Line
ADL	Advance Decline Line
ADR	Advance Decline Ratio
DIF	Difference
DEA	Difference Exponential Average
RSI	Relative Strength Index
OBV	On Balance Volume
STD	Standard
BFGS	Broyden Fletcher Goldfarb Shanno

## List of Appendices

Appendices 1 BP neural network input vector principal component score matrix	69
Appendices 2 Relevant principal component analysis	71
Appendices 3 Original data normalization	73
Appendices 4 Cwfac module	75
Appendices 5 Cwprint module	77
Appendices 6 Gabpnet program	78
Appendices 7 gabpEval program	80

## Chapter 1

### Introduction: The main research contents and thesis structure

The main research content of this report is the BPNN application in the short-term stock price prediction. By trying to break through the bottleneck of BP algorithm<sup>[1]</sup> in the short-term stock price forecast - dimension reduction problem, problem of low prediction accuracy and easy to fall into local minimum point problem. As the stock market are expanding, a large number of valuable information to mining. Therefore, in this paper, the main analysis technique combines genetic algorithm and BP neural network technology, in an integrated forecast model for the short-term stock price change carries on the preliminary discussion. In this paper, the main research contents include the following aspects:

(1) Dimension reduction methods about PCA. The BPNN to predict data should avoid the danger of "multidimensional". The NN input dimension should not be too much, otherwise it will not only reduce the learning efficiency, but also makes the prediction accuracy. In this paper, according to the characteristics of data to predict goodness-of-fit matrix to revised PCA.

(2) The optimization of BPNN. BPNN to predict data tends to lead to meet training trapped in local minimum accuracy problem. This article selects the genetic algorithm as the optimization algorithm of BPNN. On the basis of the improved BPNN prediction model is proposed and the simulation experiment in the Matlab7.0.

(3) Comprehensive model design and implementation of the short-term stock price forecast. In front of the integrated algorithm analysis and research, this article based on Matlab7.0 short-term stock price prediction model is established, at the same time, this article has carried on the empirical analysis, proves the improved algorithm combines the has certain validity, rationality and feasibility.

Surrounding these content, this paper overall architecture is as follows:

First chapter, introduction: Introduces the research background and significance of this article, the current research literature at home and abroad are reviewed, to clarify the current research status;

Second chapter, Literature survey: the stock short-term prediction of basic theory and method. Introduces the characteristics and difficulties, short-term forecasting of the stock price influence factors of affect stock price

movements. At the same time, introduces several common stock trend analysis method, and the characteristics of the various methods.

Third chapter, Theory & The Proposed Technique: the BPNN and PCA. This paper introduces the basic principle of BPNN, the advantages and disadvantages of the BPNN and the applicability of the application in the stock price forecast. At the same time, according to the principle of PCA, a combination of the stock price data to predict the actual demand, the innovation on the basis of the traditional PCA.

Fourth chapter, The Proposed Technique: genetic algorithm optimization<sup>[4]</sup> of BPNN. Firstly introduces the principle and steps of genetic algorithm<sup>[11]</sup>, through the analysis of the genetic algorithm, the improved BPNN model, and illustrates the basic steps of genetic algorithm<sup>[11]</sup> to optimize the BPNN;

Fifth chapter, Simulation results and discussions: short-term stock price forecast system establishment and implementation. First has carried on the system's overall design and detailed design, on this basis, simulation system was introduced in detail. Simulation results and discussions: the experiment and result analysis. For the proposed relative PCA and genetic neural network<sup>[13]</sup> model, has carried on the empirical analysis, respectively, by the experimental results and analysis verify the validity of the algorithm is proposed in this paper.

Finally, Conclusions and future works: briefly summarizes the full text of the main research contents innovation points, the insufficiency, and the next research direction.

## Chapter 2

### Short-term price prediction fundamental theory and method

#### 2.1 Factors of affecting stock prices

A number of factors which affect stock price<sup>[3]</sup> movements, which basically can be divided into the following categories: economic factors, business factors, political factors, industry factors, psychological factors and human factors.

(1) Economic factors.Changes in stock market,economic development and the change of the situation.The economic factors that affect the stock price with the cycle of economic growth, interest rates, exchange rates and other economic indicators.Economic operation and the stock market, the stock market is the "barometer" of economic running state, when the economy is in good condition, stock price tends to rise;While the economy in the doldrums, the stock price tends to decrease.

(2) Enterprise factors.The enterprise factors mainly include the enterprise's profit, market share and profitability of products.As well as the board of directors of enterprise management level and principal adjustment, etc.First of all, corporate earnings.Enterprise's profit level directly affect the amount of stock dividends.Stock dividends will directly cause the change of the stock supply and demand.Stock prices have a decisive role.In general, the enterprise profit level rises, the stock price will rise;the enterprise profit level fall, profitability, stock prices will fall.Both show the same change trend.Second, the profitability of product.Product profit ability, high market share will bring good advertising effect, establish a positive corporate image, and promote the growth of corporate profits, enhance the confidence of the investors, leading to the enterprise stock prices. On the other hand, will cause a drop in stock prices.

(3) Political factors.Political factors include the contents of the broad, mainly including domestic politics, international political situation, major diplomatic events, regional conflicts, etc.These factors have characteristics of large influence, sudden strong.The impact on stock market has a larger uncertainty, often makes the volatility of the stock market.

(4) Industry and regional factors. Mainly refers to the industry's market prospects for development and regional economic development level of fluctuations. Industry and regional factors is between macro and micro influence factors, its effect on stock market prices are mainly structural. First, in terms of industry, the development of every industry has its own cycles, will suffer from buds to grow and decline change development process. The change cycle called the industry life cycle. From the Angle of economics, industry life cycle can be divided into four periods, namely initial (also called juvenile stage), growth, stability and recession. In different development period of the industry there is big difference in the economic conditions and market potential. These factors will affect the choice of investors eventually be reflected in the stock price changes. Chips are the industry's share price, typically show slightly higher trend. On the contrary, in recession industry stock prices are typically shows the tendency of low gradually. Secondly, in terms of regional factors, due to the regional economic development, regional different development conditions, such as investment activity in the stock price of different areas. It will be affected by the surrounding economic environment and there are different, even in the same industry's share price will also be different. In regional economic development is good, the investment activity of the stock investment is expected, stock prices generally will be higher; On the other hand, the regional economic development lags behind, information block areas, its stock price will broadly flat down trends.

(5) Psychological factors. The social public investment groups such as psychological change has a decisive effect on stock prices. If investors are generally of a stock market prospect to disparage, regardless of the actual development status of company and selling company stock, cause a drop in stock prices. Especially the stock investors widespread in the "herd mentality", are more likely to cause the phenomenon of copycat selling stock, leading to more severe domino effect.

(6) Human factors. In the normal stock market, stock price changes are often subject to large financial controls the consortium, rather than the control of small and medium-sized investors and retail investors. For example, some large consortium to take advantage of their own money and market influence, to design a good buy and sell stocks. Prompting some stock price ups and downs in the trading market, stock by inciting shares of buy low and sell high, so as to obtain huge profits.

## 2.2 The difficulty to predict the price of stock

Stock price fluctuations by interference factors, existing difficulties mainly has the following several aspects:

(1)Uncertainty investor of stock market and buying and selling of uncertainty affects stock price forecast accurately. This influence over time and become more unpredictable.

(2)High noise data of stock price ;The influence factors of stock price, stock price data often contain more noise.The data of high noise will affect stock price prediction precision of the prediction system, also become one of the stock price prediction difficult problem.

(3) Nonlinear data of the stock price;The stock price data and the influence factors of affect stock price data shows the characteristics of highly nonlinear.Stock price forecasting system should have the powerful features of the nonlinear data processing.The traditional prediction method is the method to solve the problem of linear, however stock price prediction is a highly nonlinear complex system.

To sum up, the stock price prediction is a affected by many factors and effects of complex systems with nonlinear characteristics.

## 2.3 Stock price forecasting technology methods

Since people began to study the stock price fluctuations, they produced a variety of research methods.The most commonly used in the stock market academia analysis method mainly has the following kinds: traditional analysis method, fundamental analysis<sup>[1]</sup> and technical analysis<sup>[1]</sup> time series analysis<sup>[2]</sup> and neural network analysis<sup>[1]</sup>, gray prediction method<sup>[1]</sup>, the wavelet analysis method<sup>[1]</sup>, etc. In the face of this several kinds of methods are briefly introduce as fellows.

### 2.3.1 The traditional analysis method

Traditional analysis methods will affect the stock price factors as the research subject, through the analysis of the influence factors to realize for the forecast of stock price change trend, in order to achieve the purpose of

predicting the stock price. But the fundamental analysis and technical analysis method adopted by the technical route was different.

#### (1) Basic analysis

Basic analysis method that the world economy level of development, industry development prospects and current situation of the development of company's decision about stock value and stock decision factors, through the analysis of these factors to evaluate the fair value of the stock and, so as to realize the prediction of the stock price, form business advice accordingly. Fundamental analysis generally analysis from the following several aspects:

- 1) The level of economic development. For economic development policy, public economic indicators is analyzed, in order to infer the stock price movements.
- 2) Industry development prospects. Analysis of the stock of the industry development potential and regional economic development level on the impact of stock price change.
- 3) Company development present situation. Analysis of enterprise management system level, personnel structure and core competitiveness, etc. The influence of factors on a company's stock price.

The influence factors of the basic analysis method is difficult to quantify, it relies on long-term trace analysis and research, influenced by artificial analysis is bigger. Therefore, the basic analysis method is more suitable for long-term trend analysis of stock price, and difficult to simulated by computer.

#### (2) Technical analysis method

Technical analysis is a quantitative analysis method, mainly rely on quantitative indicators of the stock. Such as opening price, closing price and volume value. Difference of the basic analysis is the comparison about technical analysis focus on the market analysis. Technical analysis method is suitable for short-term analysis of the stock market[7], but it is difficult to predict the long-term stock price changes. Technical analysis attaches great importance to the terms of the number of changes, through to study the factors affecting amount of the change on stock market, and predict the short-term stock price changes. Technical analysis relies on three the theory premise .

- 1) Market behavior decides everything, supply and demand determine stock prices

Short-term fluctuations can cause the stock price is the most fundamental factors of supply and demand, along with the economic factors and psychological factors influencing factors, these factors will eventually reflect to

the volume and clinch a deal the price, to join a factor cannot be reflected in the stock trading volume and stock price, that this factor is irrelevant factors. To analyze the relationship between factors and stock volatility changes will need to be done by technical analysis, for example, supply exceeds demand<sup>[4]</sup>, prices will fall; demand exceeds supply<sup>[4]</sup>, prices will rise. Technical analysis focuses on market behavior will be kind of influence on the price, and do not to pay attention to what is the cause of this effect.

## 2) History will repeat itself

In past more than a century of market and technology research, some indicators and charts in the analysis more accurate response to market changes, these changes will continue in the future, based on psychological factors based on such a basis, we can predict the future by the past.

## 3) Price change trend

Market changes indeed has trend as follow: in the short term, prices follow trend, the possibility of a developing trend continues to turn is greater than the possibility of a change in trend in general. Which will continue for a period of time, barring some outside forces make it purposefully to stop or even reverse it.

### 2.3.2 Time series analysis method<sup>[5]</sup>

Time series analysis<sup>[5]</sup> is statistical method, using historical data to analyze the influence of relationship between data and change trend of future. On the basis of data changes over a period of time to predict the future. Moving average method, regression analysis, the seasonal change analysis and cyclical fluctuation analysis.

#### (1) Moving average method<sup>[5]</sup>

Moving average method<sup>[5]</sup> can be divided into simple moving average<sup>[5]</sup> and weighted moving average<sup>[5]</sup>. Moving average method through a period time count in turn get a series of temporal average, weaken the data. Moving average, however, could not overcome cycle changes brought about by the time the impact of different cycle intervals. Second, the forecasting information is not complete, moving average for the long-term trend, the sequence of ending without moving average, also can produce adverse effect to predict.

#### (2) Regression analysis<sup>[5]</sup>

Regression analysis is a commonly used statistical analysis method, this method is based on time as the independent variable, in order to form time series of statistical index as the dependent variable, using the least squares method to establish the regression relationship between the dependent and independent variables and model, on the basis of the fitting results of time series data to forecast future changes. Long-term change trend of time series contains linear trend and the trend of two curve, therefore, can according to the characteristics of the change trend, respectively, set up the corresponding linear regression model and the curve regression model.

### (3) Seasonal variation analysis<sup>[6]</sup>

Seasonal variation analysis method is used to solve the statistical indicators with seasons change and presents a certain regularity of time series problems. Seasons change analysis based on time series data whether there is a trend change and different, the most commonly used method is the quarterly average method.

### (4) Cyclical fluctuation analysis method<sup>[6]</sup>

Analysis methods mainly include periodic fluctuation trend rate and cycle remainder ratio method. If the statistic data of time series data in a certain time period along its long-term average showed a trend of certain scope and magnitude of cyclical fluctuations, the cyclical fluctuation analysis method is used to forecast.

## 2.3.3 Neural network<sup>[1]</sup> prediction method

Neural network<sup>[1]</sup> is based on a preliminary understanding of the human brain information processing mechanism and put forward a distributed parallel data processing model. Neural network according to the training data continuously adjust the relationship between the internal nodes to achieve the objective of the study data change rule, which can realize data changes over a period of time in the future trend of prediction. Neural network has the parallel processing, distributed storage, highly nonlinear, fault tolerance and good self-learning, self-organization and adaptive ability and other characteristics, market analysis and forecast in the field of economic applications. Neural network prediction methods mainly include the error back propagation (BP)<sup>[1]</sup>, radial basis function (RBF) neural network<sup>[1]</sup> and perception neural network<sup>[1]</sup>, etc.

### 2.3.4 Several forecasting methods

#### (1) Grey prediction<sup>[4]</sup> method

The Grey system theory<sup>[4]</sup> is that on the one hand contains known information, on the other hand, contains unknown, uncertain information system is analyzed when prediction is essentially to change within a certain amount of azimuth, time dependent in the process of Grey prediction. Though the analysis of the system analysis of phenomenon is random, there is no obvious regularity, but after all is orderly, bounded, so it is a collection system which has intrinsic regularity, Grey forecasting is the gray model was established based on the regularity of the gray system analysis and forecasting. Grey forecasting method to the stock market as an alternation of uncertain information system modeling, on the basis of the Grey model for analysis and forecasting of stock market.

#### (2) Wavelet analysis method<sup>[3]</sup>

Wavelet analysis is developed in the 1980 s an important analysis tool, the analysis of non-stationary data can play a good role. Wavelet analysis can more accurately extract the unstable steady trend of data sequences, which can realize the simulation of non-stationary data sequence. Due to the non-stationery and regularity of the stock market, the characteristics of the combination of wavelet analysis is analysis of the stock market is a very effective mathematical tool.

#### (3) Markov prediction method<sup>[5]</sup>

Markov method to have to deal with problems had "no aftereffect", namely when the system in the condition of known state at a certain moment, the change of system in the future only associated with the current moment, and has nothing to do with this moment before. Markov method through the establishment of the markov chain<sup>[5]</sup> stock price movements' <sup>[2]</sup> prediction, but this method is a kind of probability prediction method. The stock price prediction results can only be concluded as stock price changes in the future for a period of time within a certain numerical probability. It was unable to confirm whether the stock price will certainly to this value.

## 2.4 Summary of this chapter

This chapter first we introduces the influence factors of affect stock price. The influence factors of stock price, changes in uncertainty, more complex relationship impact on stock prices. In this case, the stock price forecast have a lot of difficulties. But these difficulties make the stock price predictions become hot spot of the research. This chapter also introduces the stock price forecast, some common methods and their characteristics, for the subsequent chapters of analysis and study laid a theoretical basis.

## Chapter 3

### Principal component analysis(PCA)<sup>[2]</sup> and BPNN<sup>[1]</sup>

Through the analysis of difficulties of stock price forecasting, factors affect the relationship between predicted values, BPNN is good at dealing with complexity and nonlinear data characteristics problems. However, the BPNN has strong sensitivity to input data measuring tool, if the input data quantity is too much, so will cause NN training time is too long, forecast data no problem, have the danger of "multidimensional". Therefore, the BPNN input data dimension reduction is about to be very important. This chapter firstly introduces the basic theory of BPNN, then introduces commonly used dimension reduction processing method about the PCA, finally to meet the needs of short-term stock price prediction and BPNN, the characteristics of innovation in PCA is put forward.

BPNN based on BP algorithm<sup>[1]</sup> for network learning, which take weights training for nonlinear differential function of multi-layer feed forward neural network. In 1986, the team of scientists led by Rumelart and McClelland multi-layer perception with nonlinear continuous transformation function has carried on the detailed analysis of error back propagation algorithm<sup>[1]</sup>, realized about multi layer networks. Due to the training of the multilayer perception often uses the error back propagation(BP) algorithm, people are also often called the multilayer perception [2] direct BP NN.

With three layers BPNN to explain the process of the derivation of mathematics, as shown in figure 3.1 for the three layers BPNN.

Network input vector  $X = (x_1, x_2, \dots, x_p)$ , looking at the output vector  $O = (o_1, o_2, \dots, o_q)$ , the input vector of the hidden layer<sup>[6]</sup> unit  $G = (g_1, g_2, \dots, g_s)$ , and the output of the hidden layer<sup>[6]</sup> unit vector  $H = (h_1, h_2, \dots, h_s)$ , and the input vector of the output layer<sup>[6]</sup>  $M = (m_1, m_2, \dots, m_q)$ , and the output of the output layer<sup>[6]</sup> vector  $Y = (y_1, y_2, \dots, y_q)$ , the output layer to the weights<sup>[6]</sup> between hidden layer<sup>[6]</sup>  $\{W_{ij}\}$ , the weights between hidden layer<sup>[6]</sup> to output layer<sup>[6]</sup>  $\{V_{jt}\}$  implied output layer threshold for  $\{\theta\}$ , output threshold for each unit of output layer<sup>[6]</sup>  $(\gamma)$ ,  $\alpha$  for the learning rate<sup>[6]</sup>.

To simulate nonlinear characteristics of biological neurons, and its corresponding function network (S) as follows:

$$f(x) = \frac{1}{1 + e^{-x}} \quad (3-1)$$

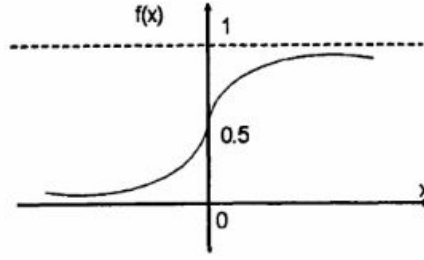


Fig 3.0 Graph of corresponding function

First use of the input vector  $X$ , the connection weights<sup>[6]</sup> of  $\{W_{ij}\}$  and threshold calculation of hidden layer<sup>[6]</sup> units  $\{g_j\}$  enter  $\{g_j\}$ , then use  $\{g_j\}$  by S function to calculate the output of hidden layer<sup>[6]</sup> of each units  $\{h_j\}$ .

$$g_j = \sum_{i=1}^n w_{ij} \cdot i_i - \theta_j \quad (3-2)$$

$$h_j = f(g_j) \quad (3-3)$$

Using the hidden layer output  $\{h_j\}$ , connection weights  $\{V_{jt}\}$  and threshold<sup>[6]</sup> calculation input output layer units  $M$   $\{\gamma\}$ , then use  $M$ , though the S to calculate the response of the output layer  $Y$ .

$$M_t = \sum_{j=1}^p v_{jt} \cdot b_j - \gamma_t \quad (3-4)$$

$$Y_t = f(M_t) \quad (3-5)$$

With the desired output vector  $O = (o_1, o_2, \dots, o_q)$ , network actual output vector  $Y = y_1, y_2, \dots, y_q$ , each unit of output layer general error  $\{d_t^k\}$ .

$$d_t^k = (o_t^k - N_t) \cdot N_t(1 - N_t) \quad (3-7)$$

With connection weights<sup>[6]</sup> of  $\{v_{jt}\}$ ,  $\{d_t^k\}$  output layer<sup>[6]</sup> of generalization error<sup>[7]</sup> and the middle layer<sup>[6]</sup> of each output  $\{h_j\}$ , calculate the middle tier units of general error<sup>[7]</sup>  $\{e_j^k\}$ .

$$e_t^k = [\sum_{t=1}^q d_t \cdot v_{jt}] h_j (1 - h_j) \quad (3-8)$$

With the generalization error of the output layer  $\{d_t^k\}$  and middle layer each output  $\{h_j\}$  (clock) fixed connection weights and threshold  $\{\gamma_t\}$

$$v_{jt}(N+1) = v_{jt}(N) + \alpha \cdot d_t^k \cdot h_j \quad (3-9)$$

$$\gamma_t(N+1) = \gamma_t(N) + \alpha \cdot d_t^k \quad (3-10)$$

In the middle layer of each unit of generalization error  $\{e_j^k\}$  and the input layer of each unit of the input vector  $X = (x_1, x_2, \dots, x_p)$  fixed connection weights of  $\{w_{ij}\}$  and threshold values  $\{\theta\}$ .

$$w_{ij}(N+1) = w_{ij}(N) + \beta \cdot e_j^k \cdot \alpha_i^k \quad (3-11)$$

$$\theta_j(N+1) = \theta_j(N) + \beta \cdot e_j^k \quad (3-12)$$

Repeated the given mode, until BP neural network<sup>[1]</sup> convergence within the given scope of permissible error value or set the number of training. fig3.1

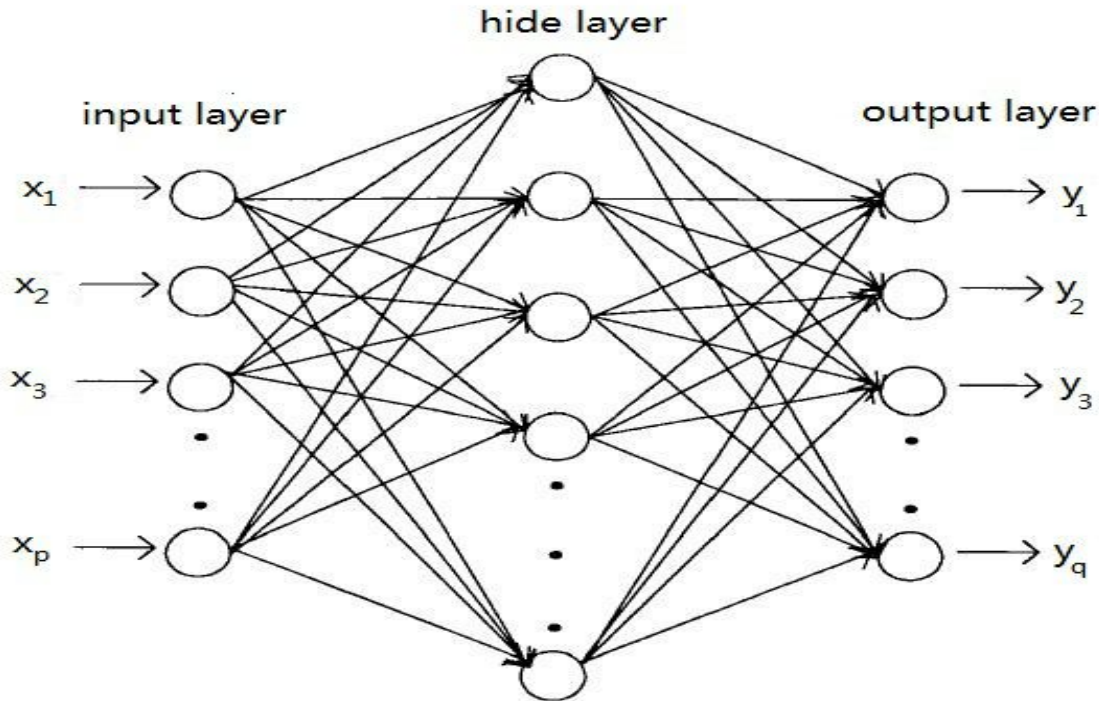


Fig3.1 Three layer BPNN

### 3.1 The characteristics of the BPNN

#### (1) Advantage of BPNN

1) Nonlinear mapping function. BPNN to establish a nonlinear function relationship between input and output data, the BP algorithm<sup>[1]</sup> of mathematical derivation proves that contains only one hidden layer<sup>[6]</sup> of BPNN can approximate any nonlinear continuous function<sup>[2]</sup> at arbitrary precision. Therefore, the BPNN can deal with relationship between the internal fuzzy prediction problem or more complex system.

2) Self learning and adaptive functions<sup>[7]</sup>. The BPNN in the process of training data, can be implemented through continuously adjustment of the weights<sup>[6]</sup> of the input data and output data between implicit rules of learning, at the same time, adaptive to learning content cure above the weights<sup>[6]</sup> of the BPNN.

3) Generalization ability<sup>[7]</sup>. BPNN generalization ability refers to the design of pattern classifier, on the one hand, make sure the BP into network can to correct the problems need to be solved by classification, on the other hand also make sure that the BP after data into the network learning and training, the ability to rule or not trained had to identify the noise pollution data accurately.

4) Fault tolerance. To some individual neurons in structure of BP neural network<sup>[1]</sup> to forecast the damage, and the predicted results will not be too big impact. It shows that BP network can be partially damaged without affecting its use effect. Therefore, BP neural network<sup>[1]</sup>'s input data has a strong ability of fault tolerance of noise.

## (2) Defects of BPNN.

BP algorithm<sup>[1]</sup> was applied to three layer perception with nonlinear transformation function[5], which can approximate any nonlinear function<sup>[5]</sup> with arbitrary precision. The advantage of BPNN is used widely. However, standard BP algorithm<sup>[1]</sup> in the application also exposed the inadequacy of several aspects:

1) Easy to fall into local minimum point. In the process of optimization to obtain the global minimum point Traditional BP algorithm is a kind of local optimization algorithm, the BP neural network to deal with the problem is often internal influence mechanism of complex nonlinear problems. Traditional BP algorithm<sup>[1]</sup> by local optimal direction gradually to adjust and optimize the network parameter values when the parameter value to local minimum point, BP algorithm cannot be achieved from the local minimum point to continue to search optimization function, cause the network parameters converge to local minimum point can not escape. In this case, BPNN weights<sup>[6]</sup> and threshold of the initial value of NN training effect using disparities initial values often lead to network convergence in disparities local minimum points. Which leads the network training for many times but get the big difference of result.

## 2) BP algorithm optimization way leads to low learning efficiency

The BPNN optimization problem has certain complexity, however the weights<sup>[6]</sup> and threshold<sup>[6]</sup> of BPNN optimization method using the gradient descent algorithm. Therefore, BP algorithm<sup>[1]</sup> in the process of getting the optimal solution will not only appears when close to the minimum point<sup>[8]</sup> optimization, the speed will slow down. The phenomenon of decline is reduces the learning speed of BP algorithm; At the same time, to perform in the BPNN model first update method of gradient descent step value shall be assigned to the BP NN. It can't use the traditional one dimensional search method for each iteration step length. In this way BP algorithm<sup>[1]</sup> will affect the efficiency of optimization. The above two aspects to BPNN algorithm in slow convergence speed of faults appeared in the process of training.

## 3) Determine the network structure of a lack of authority theory instruction

Selection of structure parameters of BP neural network is generally determined according to experience, the main answer is trial and error method<sup>[6]</sup>, namely, by using the method of training for many times to take the optimal current there is no authoritative method, which caused the BP neural network tend to have larger redundancy, lowering the efficiency of online learning.

### 3.2 BPNN model in the stock price advantages and problems in short-term prediction<sup>[8]</sup>

The stock market<sup>[4]</sup> has characteristics of large amount of data and complex factors. At the same time, it can cause stock market volatility factors and has the characteristics of diversity and complexity. As above mentioned in section, the BPNN has the nonlinear simulation, the advantages of learning itself, in the financial markets by the broad masses of data analysis in the favor of the researchers, and has achieved some fruitful research results. The basic process of general BPNN prediction model as shown in Fig 3.2.

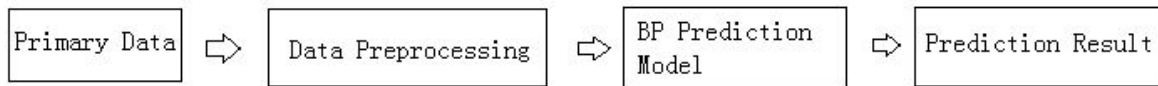


Fig 3.2 The basic structural diagram of the BP prediction model<sup>[1]</sup>

With the deepening of the BPNN research, it has some defects are gradually exposed. Specific to short-term stock price<sup>[8]</sup> forecast, basically has the following two points:

#### (1) Input data dimension reduction about the problem

The original data is mainly can be divided into technical analysis indicators and stock price data. Technical analysis indicators commonly used include stocks on the day of the opening price<sup>[9]</sup>, the highest price<sup>[9]</sup>, closing price<sup>[9]</sup>, the lowest price<sup>[9]</sup>, volume<sup>[9]</sup>, etc. At the same time, there are calculated based on the experience of various economic theories and technical indicators. If all as the input vector of the prediction model, that is unbearable for forecasting model. Therefore, how to choose from many of the technical indicators of valuable index as input vector and how will these indicators to forecast model input is to build a predictive model have been the important problem.

#### (2) Inherent defect of BP neural network<sup>[1]</sup> to overcome the problem

The intrinsic defects of BP algorithm in the previous section 3.1.2 has done a brief narrative. How to combine the characteristics of the stock price data to optimize the BP neural network<sup>[1]</sup> model, to overcome the easy to fall into local minimum point and affected the accuracy of model prediction problem is also the research contents of this paper.

To sum up, using the BP neural network<sup>[1]</sup> to forecast the stock price to the dimension of input data about problem and local minimum point.

### 3.3 Dimension about the BPNN

Due to the generalization ability of the BP neural network<sup>[1]</sup> self-organizing, adaptive characteristics, are widely used in artificial intelligence, pattern recognition and data forecasting. In the field of stock price forecasting, the price of stock data are comprehensive. However, there are many influence factors of stock price and the number of various factors affecting data quantity is big. Strong correlation characteristics, often appears redundant input data<sup>[8]</sup> of NN. It is not only reduces the efficiency of NN learning, but also influence the predicted after the training. Which requires a large number of data from the original as far as possible which contain more effective information about reducing data appropriately. By using the processed data as input data of BPNN. Therefore, the purpose of this article studies is use stock price data preprocessing mainly concentrated in the input data dimension reduction about aspects.

### 3.4 Commonly used method of dimension reduction of the commonly used method

Commonly used dimension reduction method is about self-organizing feature mapping method (SOM) and PCA. Self-organizing feature map (SOM)<sup>[5]</sup> is a kind of non-teacher learning method and self-organizing feature mapping model includes local interconnection neuron array, comparative selection mechanism, and an adaptive process. Process neuron array is related to input from the solution space, at the same time, these simple discrimination function of the input signal. Local interconnection function at the same time inspire the selected processing unit and the neighboring connection neurons, the adaptive process correction inspired the parameters of processing units, thereby increasing its discrimination function corresponding to specific input

output values. This is a kind of neural network to draw from the input signal and its important feature of the method. SOM method is a kind of non-teacher learning method, this method there is a long training time, and hard to explain.

PCA is solving the problem of dimension reduction about a kind of effective method, also is a kind of commonly used methods. PCA is a kind of multivariate statistical analysis, by constructing co-variance matrix, to subtract matrix eigenvalue and eigen-vector method. this paper selects PCA method of the reduction as the data dimension, the following a brief introduction to the basic principle of PCA.

### 3.5 Introduction of principle of principal component analysis<sup>[2]</sup> (PCA)

PCA is a model that get original data vector back together as a set of independent each other less several comprehensive index vector to replace the original index vector. At the same time, to reflect the original data contains the main information. The mathematical principle of PCA are briefly introduced as follows<sup>[8]</sup>

$$X = (x_1, x_2, \dots, x_p) = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \quad x_j = \begin{pmatrix} x_{1j} \\ x_{2j} \\ \vdots \\ x_{nj} \end{pmatrix} \quad j=1, 2, \dots, p \quad (3-13)$$

P observation variable transform into p a orthogonal vector:

$$\begin{cases} A_1 = a_{11}x_1 + a_{12}x_2 + \cdots + a_{1p}x_p \\ A_2 = a_{21}x_1 + a_{22}x_2 + \cdots + a_{2p}x_p \\ \quad \quad \quad \dots \\ A_p = a_{p1}x_1 + a_{p2}x_2 + \cdots + a_{pp}x_p \end{cases} \quad (3-14) \quad [11]$$

Calculation shall meet the following conditions:

(1) The  $A_i$ , the  $A_j$  are unrelated ( $i, j = 1, 2, \dots, p$ ).

(2) Meet  $A_1$  Each vector variance, behind the variance is greater than its variance is greater than its behind each to  $A_2$  Quantity variance, ...,  $A_{p-1}$  is greater than the variance of  $A_p$

(3) Meet  $a_{k1}^2 + a_{k2}^2 + \cdots + a_{kp}^2 = 1$ , after  $k = 1, 2, \dots, p$

Here, the definition of  $A_1$  as the first principal component,  $A_2$  as the second principal component,... And therefore, results total P principal components can be obtained. The type, the skunk said coefficient of main component.

Calculation process in the form of a matrix can be expressed as:

$$A = \begin{pmatrix} A_1 \\ A_2 \\ \vdots \\ A_p \end{pmatrix} \quad X = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_p \end{pmatrix} \quad a = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_p \end{pmatrix} = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1p} \\ a_{21} & a_{22} & \cdots & a_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ a_{p1} & a_{p2} & \cdots & a_{pp} \end{pmatrix} \quad (3-15)$$

Principal component matrix can be expressed as:  $A = aX$

### 3.6 Relevant principle of PCA<sup>[11]</sup>

PCA is a kind of dimension technology, which mainly multiple factors affecting transformation and combination of less analysis method, the factors affecting the several influence factors include the influence factor of the original information as much as possible. But in terms of analysis of factors affecting the stock price, PCA has certain blindness in extracting principal component algorithm simply consider the various factors affecting on principal component of load, without considering the influence factors and correlation analysis target. Two vectors in the econometric asked the goodness-of-fit of the correlation of two variables can be described quantitatively the size, in order to make up for the defect of principal component analysis in this aspect, this article consider joining the influencing factors and analysis of target goodness-of-fit matrix, based on the goodness of fit matrix to revise correlation matrix, thus put forward a new PCA, we define it as the main component analysis method. Related algorithm specific principle of PCA is introduced as follows:

Suppose there are n samples, each sample has p variables, constitute the data matrix:  $n * p$  :

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix} \quad (3-16)$$

(1) computing the correlation coefficient matrix<sup>[11]</sup>

$$M = \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{12} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \vdots & r_{pp} \end{bmatrix} \quad (3-17)$$

$r_{ij}$  (i, j = 1, 2,..., p) for the correlation coefficient<sup>[12]</sup> of the original variable  $X_i$  and  $X_j$ ,  $r_{ij} = r_{ji}$ , its computation formula is:

$$r_{ij} = \frac{\sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j)}{\sqrt{\sum_{k=1}^n (x_{ki} - \bar{x}_i)^2 \sum_{k=1}^n (x_{kj} - \bar{x}_j)^2}} \quad (3-18) \quad [11]$$

(2) Calculation of influencing factors and the target variable of goodness of fit<sup>[11]</sup>

It can be introduced the concept of goodness of fit in econometric to represent the target variable and the factors affecting the relationship. Goodness-of-fit  $R^2$  can use Venn visual representation is as follows:

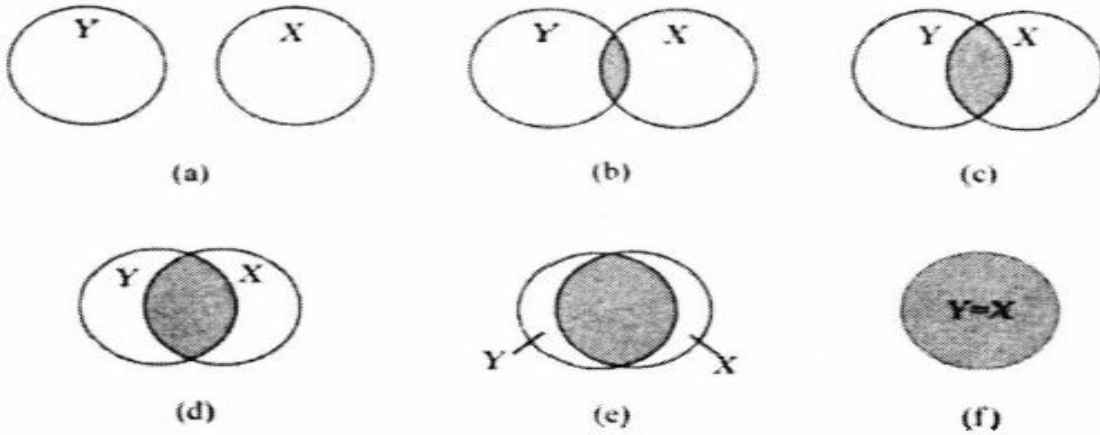


Fig 3.3 Venn diagram of good of fit

In the diagram above, Y is the target variable, X is the influencing factors of variable, goodness-of-fit<sup>[11]</sup>  $R^2$  degree of explanation for variables X to Y, the more overlapping part shows that the greater the degree of interpretation, the greater the corresponding correlation. Goodness-of-fit<sup>[11]</sup>  $R^2$  calculation process is as follows:

Using the least squares fitting variable<sup>[9]</sup>  $x_i$  and  $Y_i$ , estimated equation is obtained:

$$Y_i = \hat{\beta}_0 + \hat{\beta}_1 X_i + \hat{u}_i \quad (3-19)$$

$\hat{\beta}_0$  and  $\hat{\beta}_1$  is estimated parameters, the  $\hat{u}_i$  is residual items. Equation can be expressed as

$$Y_i = \hat{Y}_i + \hat{u}_i \quad (3-20)$$

And

$$R^2 = \frac{\sum (\hat{Y}_i - \bar{Y})^2}{\sum (Y_i - \bar{Y})^2} \quad (3-21)$$

Let

$$R = \begin{bmatrix} R_1^2 & \cdots & 0 \\ \vdots & R_2^2 & \vdots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & R_p^2 \end{bmatrix} \quad (3-22)$$

(3) Calculate correlation matrix<sup>[12]</sup>

Use goodness-of-fit matrix correction correlation matrix

$$M_1 = RMR = \begin{bmatrix} R_1^2 & \cdots & 0 \\ \vdots & R_2^2 & \vdots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & R_p^2 \end{bmatrix} \begin{bmatrix} r_{11} & r_{12} & \cdots & r_{1p} \\ r_{21} & r_{22} & \cdots & r_{2p} \\ \vdots & \vdots & & \vdots \\ r_{p1} & r_{p2} & \cdots & r_{pp} \end{bmatrix} \begin{bmatrix} R_1^2 & \cdots & 0 \\ \vdots & R_2^2 & \vdots \\ \vdots & \vdots & \vdots \\ 0 & \cdots & R_p^2 \end{bmatrix} \quad (2-23)$$

(4) Calculate the adjusted correlation matrix eigenvalue and eigen-vector of  $M_1$

Characteristic equation of, in  $|\lambda I - M_1| = 0$  order of size  $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_p \geq 0$ , then respectively to find the eigenvalues corresponding to the feature vector.

(5) Computing ratio and total ratio

Ratio:

$$\frac{\lambda_i}{\sum_{k=1}^p \lambda_k} \quad (i=1,2,\dots,p) \quad (3-24)$$

Total ratio:

$$\frac{\sum_{k=1}^i \lambda_k}{\sum_{k=1}^p \lambda_k} \quad (i=1,2,\dots,p) \quad (3-25)$$

Generally take the eigenvalues of the cumulative contribution rate of more than 85% of the 1, 2,... And the m ( $m \leq p$ ) principal components.

(6) Principal components load calculation:

$$l_{ij} = p(z_i, x_j) = \sqrt{\lambda_i} e_{ij} \quad (i=1,2,\dots,p) \quad (3-27)$$

(7) All landowners calculation<sup>[10]</sup> of each principal component scores get the input vector<sup>[6]</sup> of BP neural network<sup>[1]</sup>

The input vector matrix projection principal component load based on the matrix of neural network input vector.

$$\begin{cases} z_1 = l_{11}x_1 + l_{12}x_2 + \dots + l_{1p}x_p \\ z_2 = l_{21}x_1 + l_{22}x_2 + \dots + l_{2p}x_p \\ \dots\dots\dots \\ z_m = l_{m1}x_1 + l_{m2}x_2 + \dots + l_{mp}x_p \end{cases} \quad (3-28) \quad [12]$$

Principal component scores are obtained:

$$Z = \begin{bmatrix} z_{11} & z_{12} & \dots & z_{1m} \\ z_{21} & z_{22} & \dots & z_{2m} \\ \vdots & \vdots & & \vdots \\ z_{n1} & z_{n2} & \dots & z_{nm} \end{bmatrix} \quad (3-29) \quad [12]$$

### 3.7 Summary of this chapter

Dimension reduction about BP neural network<sup>[1]</sup> is applied to short-term stock price<sup>[5]</sup> forecast have some important step:to build the first step in the short-term stock price forecast model.This chapter use the existing per-treatment method,comprehensive characteristics of short-term forecasting of the stock price and the characteristics of BP neural network<sup>[1]</sup>, a PCA<sup>[11]</sup> as the data prepossessing method.The final prediction accuracy as the guide, with the target and the factors affecting the vector q matrix as the correlation matrix of the goodness-of-fit<sup>[11]</sup> of value to revise the traditional PCA. To strengthen PCA principal component with the target of the correlation between vector-value for subsequent work to be done in further research of the prediction model.

## Chapter 4

### Optimize BP NN<sup>[1]</sup> model based on the GA<sup>[11]</sup>

This chapter according to the BP neural network<sup>[1]</sup> model in the short-term stock price forecast reflects the easy to fall into local minimum points, to the limitation of the global search, using the genetic algorithm<sup>[11]</sup> to optimize the BP neural network<sup>[1]</sup>'s initial weights<sup>[6]</sup> and threshold<sup>[6]</sup> to a new genetic neural network<sup>[11]</sup> prediction model.

#### 4.1 Concept and principle of GA<sup>[11]</sup>

In the 1980 s, Holland professor at the university of Michigan and the students in the study of artificial adaptive systems and the nature of puts forward the idea of genetic algorithm for the first time. GA is a kind of direct search algorithm<sup>[12]</sup>, its optimization for specific problems solution does not depend on the problem of genetic algorithm principle ideas derived from the principle of Darwin's theory<sup>[13]</sup> of evolution and genetics. At present, the genetic algorithm is widely used in artificial intelligence, social science, and other research areas.

##### 4.1.1 Basic concept of GA<sup>[11]</sup>

GA is combined with Darwin's theory<sup>[13]</sup> of evolution and genetics principle and the direct search optimization method. Therefore, the genetic algorithm in reference concept of a lot of evolution and genetics. The genetic algorithm<sup>[11]</sup> is used in many concepts in biology. This chapter will describe and explain the use of these concepts. Concept of it is worth noting that used here are different from the concept of biological.

##### (1) Gene chain code<sup>[12]</sup>

Biological character is determined by the biological genetic chain code<sup>[12]</sup>. When using the genetic algorithm, need every solution of the problem encoding become a gene chain code<sup>[12]</sup>. Assuming that the integer 1552 is a solution of the problem, we can use binary form 1552 to 1100001000 to represent the solution of the

corresponding gene chain code <sup>[12]</sup>, a representative each one gene. Therefore, a gene chain code <sup>[12]</sup> is a solution of the problem, each gene chain code <sup>[12]</sup> is sometimes described as an individual. In the genetic algorithm <sup>[11]</sup> is also sometimes called gene chain code <sup>[12]</sup> chromosomes.

## (2) Gene

Genes are the elements in the list, is used to represent the individual characteristics. A string  $S = 1011$ , for example, is one of the 1,0,1,1 this four elements are called genes.

## (3) Gene position

Gene position is a position in the string. Gene position calculated by the list of left to right, for example in the string  $S = 1101$ , 0's gene position is 3.

## (4) Group

Some individuals form group. Group is a collection of several individuals <sup>[12]</sup>. Because each individual represents a solution, so the group is a collection of the solutions. For example,  $p_1 = \{x_1, x_2, \dots, x_{100}\}$  is composed of 100 solution group.

## (5) Cross

Many of the reproduction of the organism was done by chromosomal crossover. In the use of the concept of the genetic algorithm, and the cross as a operator. The implementation process of the operator as follows: select group of two individuals  $x_1, x_2$  for both the individual gene chain code cross, parents do the produce of two new individual  $x'_1, x'_2$ , as their progeny. Simple crossover method: random selection of a cut-off point, the  $x_1, x_2$  gene chain code in the truncation point cut, and exchange by half, which combined into two new individual  $x'_1, x'_2$  (as shown in figure 4.0). In many applications, the crossover operator is based on a certain probability, the probability is called crossover probability.

	Parent			Progeny	
$x_1$	1000	10011110	$x'_1$	1000	11000110
$x_2$	0110	11000110	$x'_2$	0110	10011110

Fig 4.0 New individual combination

#### (6)Variation

Variation in the process of the reproduction of the organism is a important step.By mutations in certain genes on the chromosome location makes new individuals are different from other individuals.Mutation in the genetic algorithm is an important operation.The implementation method is as follows: the operator for an individual in the group, chain code that gene, randomly selected from a particular, namely a particular gene.Will turn the gene (0 to 1, 1 to 0) as shown in figure 4.1.


100011000110→100011010110  


Fig 4.1 Gene code chain variation

#### (7)Fitness

The degree of organisms to the environment to adapt to different show different life thereby.Here, each individual corresponds to a solution of optimization problem”  $x_1$  ”, each  $x_1$  solution corresponding to a function value  $f_i$  , the greater  $f_i$  indicates the better  $x_i$  we get, that means the higher fitness of environment.So can use the function value of each individual  $f_i$  as its fitness to the environment.

#### (6) Selection

The purpose of selection is to select from instant group of excellent individuals and make them have the opportunity to individual as their parent generation to produce offspring. Individual judgment criteria is good or not their fitness values. Apparently this was borrowed from the principle of such an evolution, namely the individual fitness is higher, the more it is a choice. As a kind of operator, select operator in the genetic algorithm has many implementations, one of the most simple way is and the fitness value is proportional to the probability method is used to choose. Specifically, is to first calculate the sum of all individual fitness in  $\sum f_j$ , then calculate the goodness of fit of each individual occupies the proportion of  $\frac{f_i}{\sum f_j}$ , and as the corresponding choice probability. In this way, the genetic algorithm<sup>[11]</sup> can find the largest fitness of individual.

#### 4.1.2 Principle of genetic algorithm<sup>[11]</sup> and basic operation

GA is a kind of widely used optimization algorithm. The principle of this algorithm is from the principles of genetics and evolution principle. Algorithm will stay will converted into the initial population as operation object. The evolutionary mechanism of "survival of the fittest", by using the method of selection, crossover and mutation of initial population genetic operation to get a new species. The new species include the original information, and is better than the last generation of population. Through continuous operation until the fitness or genetic algebra is achieved, so as to seek the optimal solution of original problem.

GA optimization process of a problem is totally random, but the effect of the genetic algorithm operation is not random. GA can use past information to effectively speculates that the next generation of information to complete the generational evolution, so as to realize the solution of the problem. Steps of GA includes following aspects: problems coding<sup>[13]</sup>, initial population initialization<sup>[13]</sup>, the design of fitness function<sup>[12]</sup>, genetic operation design<sup>[11]</sup>, the selection of control parameters, the setting of the termination of the loop condition.

##### (1) Problem encoding

The first step of genetic algorithm first converts to solving the problem of genetic algorithm to process and operation of the initial population, initial population is composed of chromosomes. This process is to solve the problem and the process of establishing corresponding relationship between chromosome. To implement this process will be selection populations of coding and decoding way. The commonly used encoding binary encoding and real encoding.

## (2) Population initialization

Action object is to solve the problem of GA coding and initial population, composed of population individuals randomly generated in general. Population scale should not be too big or too small, although population size is too large to some extent, can improve the effect of genetic algorithm to optimize, the corresponding computational efficiency but also reduces; And the effect of population size is too small will not optimize. Therefore, under the condition of don't know much about solving the problem, because don't know the number of the optimal solution, so generally adopt uniform in the problem space, randomly generated individuals initial population.

## (3) Fitness function<sup>[11]</sup>

Fitness function according to the individual survival environment, the adaptive value calculated by the fitness function, GA to select individual genetic probability. In the actual operation, the objective function are often used as the fitness function of genetic algorithm.

## (4) Genetic operations<sup>[12]</sup>

Including selection, crossover and mutation genetic algorithm (ga) three kinds of genetic operation mode.

### 1. Selection

GA using selection operator to achieve superior slightly tide of individuals in the group operation: high fitness individuals are copied to the next generation of probability; On the other hand, was tiny probability of next

generation heredity to group. Select operation task is according to the select few individuals in some way from the parent population genetic to the next generation.

The common method to several options are:

<1>The roulette wheel selection method

The basic idea: the individuals while the probability proportional to size of fitness. Specific operation is as follows:

- (1) Calculate each individual fitness value (decoding)  $f_i, i=1, \dots, M$ . (M is population size.)
- (2) Using the proportion of formula selection operator, calculate the probability of each individual is heredity to next generation;
- (3) Calculate the accumulation probability of each individual;
- (4) In  $[0, 1]$  interval to produce a uniform distribution of pseudo-random number  $\gamma$
- (5) If  $\gamma \leq q_1$ , then select the individual 1; otherwise, select individual k, makes:  $q_{k-1} \leq \gamma \leq q_k$  acuity were established;
- (6) Repeat (4), (5) M times.

Because this method is based on the probability choice, statistical error. Combined with the optimal preservation strategy, therefore, to ensure that the fitness of the best individual is able to evolve to the next generation, not destroyed by the randomness of genetic operations, guarantee the convergence of the algorithm.

<2>The expected value method

In the roulette wheel selection algorithm, when the individual is not too much, on the basis of the random number to choose may not correctly reflect the individual fitness, high fitness individuals might be eliminated, the low fitness individuals are likely to be a choice. Using the expected value method can solve the problem.

Using this method, firstly calculates the expected number of every individual to survive in the next generation of  $M_i$ .

$$M_i = \frac{f_i}{f} = \frac{f_i}{\sum f_i / N} \quad (4-1)$$

One of the  $f_i$  is the  $i^{th}$  individual fitness value<sup>[13]</sup>,  $\bar{f}$  is the average adaptive value<sup>[13]</sup>, N is the total number of individuals that population size. If an individual is selected, and participate cross. It will subtract 0.5 survival prospects; If not participate cross, the number of individual survival prospects minus 1. In both cases, if the survival of the individual expected number less than zero, then the individual choice will not be involved.

### < 3 > Sort option

This method is the main characteristic is to the individual fitness whether take positive or negative, and the numerical difference between the individual fitness, but the roulette wheel selection requirements to adapt to the value in general is a negative number. The basic idea: to all in the group according to their fitness size sorting, according to the order while the probability of the distribution of each individual. Specific operation is as follows:

- (1) List the group according to their fitness size sort;
- (2) According to the concrete solving problem, design a probability distribution table, arrange the probability value according to the above order is assigned to each individual;
- (3) Based on the probability of each individual values as a basis of inheritance to the next generation probability according to the given probability of the individuals with roulette wheel selection method to produce the next generation.

### < 4 > Best individual preservation

The basic idea is: the best individual in the parent group directly into among children. This method can ensure that the resulting in the process of genetic individuals will not be damaged by crossover and mutation operators, it is the genetic algorithm convergence is an important guarantee conditions. It is also easy to make local optimal individual is not easy to be eliminated, so that the algorithm is global search ability stronger. Combined with other coordinate selection methods used, can get satisfactory result.

### < 5 > Select random league competition

Basic idea: every time selection of N individuals with the highest fitness among individual genetics to the next generation. Specific operation is as follows:

- (1) Randomly selected from the group N individual fitness size comparison, will be one of the highest degree of individual genetic into the next generation of group;
- (2) Repeat the above process M (for the group size), can be the next generation.

## 2.Crossover

Crossover refers to two overlapping chromosomes to exchange some of its genes in some manner<sup>[14]</sup>, thus forming two new individual. The main method to generate new individual. It determines the global search ability of GA, plays a key role in genetic algorithm<sup>[11]</sup>. Here are several kinds of commonly used for binary coding or crossover operator of real number encoding.

The cross method commonly used are:

### <1>Single point of crossover

Random set a intersection in the individual string, before or after the point at which a part of the structure of the two individuals swaps, generate two new individual.

### <2>Two points crossover

Refers to the mutual matching of two individual coding string random set intersection, and exchange between the two intersection part of genes, as shown in figure 4.2.

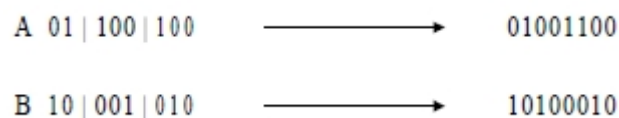


Fig 4.2 Two points crossover

### <3>Multi-point crossover<sup>[14]</sup>

Multi-point crossover<sup>[14]</sup> is an extension of the former two cross multiple intersection in the two strings.

### <4>Uniform cross

Uniform crossover is by setting up a shield, to decide on each new individual genetic inheritance which old individual genes.

As shown in figure 4.3, when the word block to 0, new individual A's gene is from old individual A; when the block word is 1, the new individual B's gene is from old individual B'.

Block	01000101		
A	00101101	A'	01101100
B	11101100	B'	10101101

Fig 4.3 Uniform crossover

#### <5>Arithmetic crossover

It is to point to by a linear combination of the two individuals and to produce new individual, arithmetic crossover is suitable for real number encoding. Between two individuals in the  $x_1, x_2$  arithmetic crossover, the cross of two new individuals  $x'_1, x'_2$  for:

$$\begin{cases} x'_1 = \alpha x_1 + (1 - \alpha)x_2 \\ x'_2 = \alpha x_2 + (1 - \alpha)x_1 \end{cases} \quad (4-2)$$

If the  $\alpha$  for interleaving, constant between  $\alpha$  to (0, 1), then this crossover operation for uniform arithmetic crossover; If the  $\alpha$  is a variable determined by the evolution algebra, then this crossover operation for non-uniform arithmetic crossover.

### 3. Mutation

Refers to the individual mutation of certain genes on chromosome coding string value with the other loci alleles to replace, thus forming a new individual. It is to generate new individual auxiliary method, it determines the local search ability of GA. When the crossover operator and mutation operator of mutual cooperation, common to complete the search space of global search and local search, so as to make the search performance of GA in good with the optimization process of optimization problems.

The use of mutation operator in GA is aimed to improve the local search ability of GA and maintain population diversity, prevent premature phenomenon.

Mutation in the commonly used several methods are:

#### <1>Basic mutation

For individual coding string mutation probability  $p_c$  randomly assigned one or several gene mutation.

#### <2>Uniform mutation

Respectively in accordance with a range of uniformly distributed random Numbers, with a smaller probability to replace individual coding string all the original gene loci on value. Uniform mutation is especially suitable for application in the early stages of the GA run phase, it makes the search point can move freely over the whole search space, which can increase the diversity of the group, makes the algorithm more patterns.

#### <3>Dual mutation

Need two chromosomes, and through the binary mutation generated after two new individuals in each gene original chromosomes were taken corresponding value with and/xor. It has changed the traditional way of mutation, effectively overcome the premature convergence, improves the genetic algorithm to optimize the speed.

#### <4>Gaussian mutation

In variable with a mean  $\mu$ , variance a  $\sigma^2$  to normal distribution to a random number to replace the original genetic value, operation are similar and uniform mutation.

### (5) Control parameter selection<sup>[12]</sup>

GA needs to determine the parameters of the main operation including group number (N) ( $p_c$ ), crossover probability and mutation probability ( $p_m$ ).

#### 1) Group number (N)

Population is genetic algorithm operation object, enough populations is the prerequisite of genetic algorithm for effective operation, ensures that the genetic algorithm to solve the effect. However, if the population is too large will cause computational efficiency is lower. In the actual operation of the populations are between 20 to 200.

#### 2) Crossover probability ( $p_c$ )

Crossover probability rules used in the genetic operations cross means the number of times, the higher the value of crossover probability, the production and the change of the individuals in the group, the more active. At the same time, the lower the crossover probability of individual and change more slowly, even may cause slow and stagnation of evolution. In practice the crossover probability values are between 0.6 and 1.00.

### 3) Mutation probability ( $p_m$ )

Mutation is a powerful tool to overcome premature convergence of genetic algorithm. The proper mutation probability can overcome the defect of premature loss of information, some gene position to achieve individual diversity of the population. But if the mutation probability value is too large will search into random search. In practice the mutation probability values are between 0.0005 and 0.01.

### (6) End loop condition

The search of genetic algorithm with a certain randomness, to avoid the infinite search and circulation loop termination conditions need to set algorithm. In practice can be the latest group fitness as a condition of termination of the judgment of the chromosomes, the fitness to reach a set range, acceptable algorithm of effect, think convergence; Or set a specific genetic algebra, when the genetic algorithm to achieve a specific genetic algebra operation is terminated when the genetic operation, the final population as the optimal solution is obtained.

## 4.1.3 Process of genetic algorithm

The basic process and structure of traditional genetic algorithm as shown in figure 4.4 and 4.5

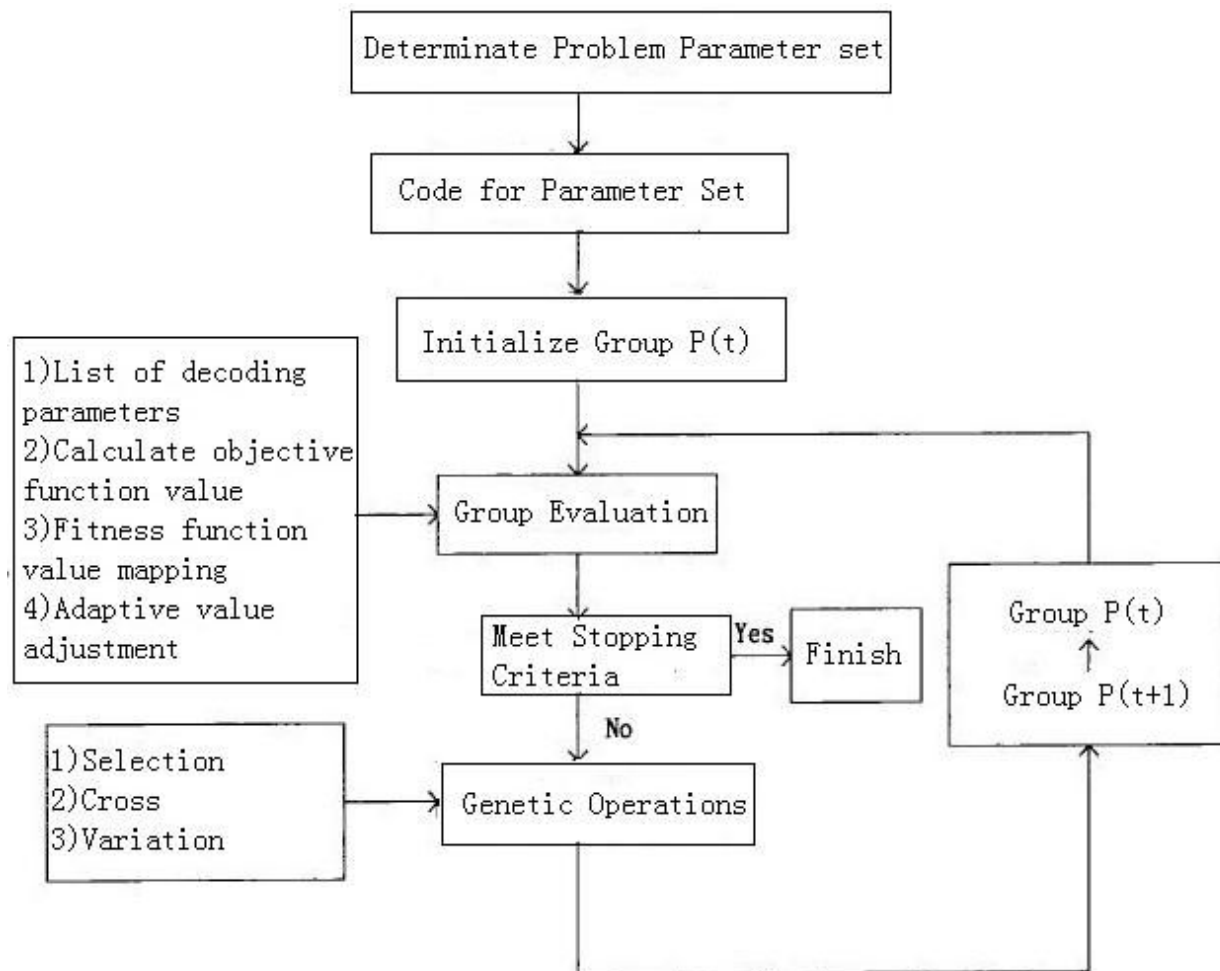


Fig 4.4 Genetic algorithm flowchart

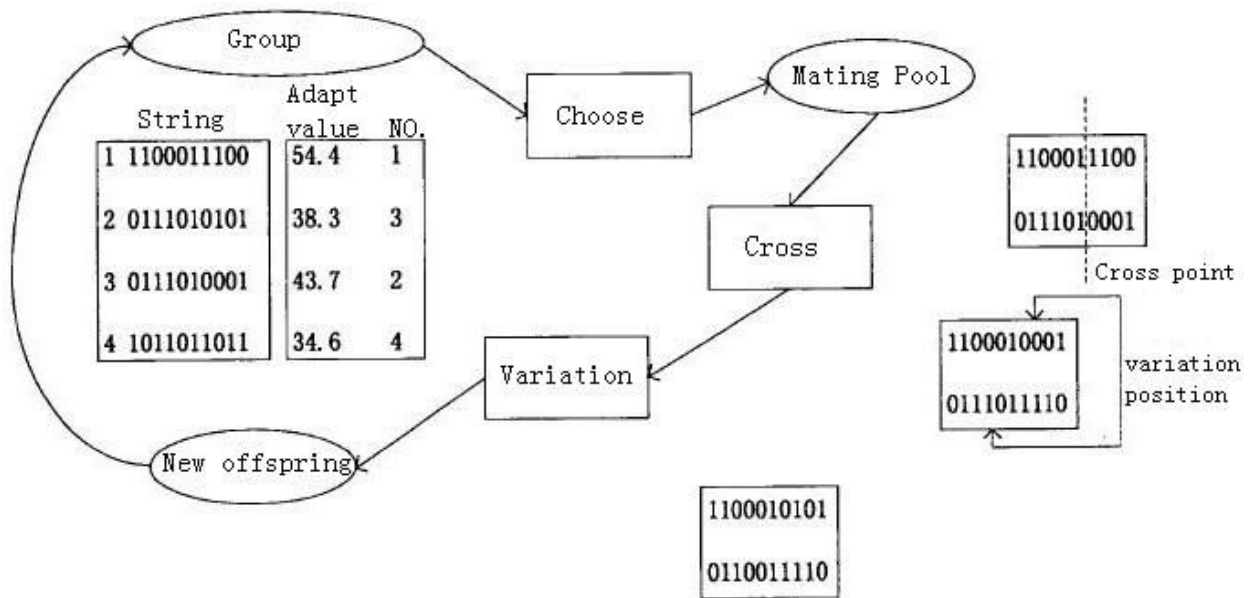


Fig 4.5 Genetic algorithm flowchart

As can be seen from the figure 4.4 and figure 4.5, the process of genetic algorithm is a process of circulation, algorithm is indispensable basic steps are as follows:

- (1) coding strategy choice. Choose the appropriate coding strategy, treat to solve the problems were analyzed, and its through the selected encoding strategy is converted into a string of structure space;
- (2) define the fitness function  $f(x)$ ;
- (3) Select the genetic strategy. To determine the initial population, the number of genetic operations and probability value of genetic parameters such as
- (4) initialization population:
- (5) computing to adapt to the value of  $f(X)$ ;
- (6) According to the genetic strategy, to genetic operation, the group of population evolution;
- (7) All landowners to determine whether a group evolution degree, to meet the set conditions, go back to step 6 does not meet the conditions, or modify the genetic strategy returns to step 6.

#### 4.1.4 The advantages of genetic algorithm

GA[8] in the treatment of the complex system optimization solution has a strong applicability, and is different from the general system optimization search algorithm, genetic algorithm processing object is a unified coding solution space, therefore, has a lot of genetic algorithm in the system optimization does not have the superiority of other algorithms, mainly reflected in several aspects:

(1) Process of GA coding space object is to solve the problem. Genetic algorithm to solve the problem of a particular form of coding as treatment object, the difference and common algorithm deal directly with the way to solve the problem. GA of chromosomes with Darwin's genetic and gene as the foundation, the concept of the genetic and variation of the introduced features into optimization algorithm, to find the optimal solution of problem solution process analogy for gene in the nature of heredity and variation process, the optimization way of no quantitative or are hard to quantify the numerical optimization has a strong adaptability, the problem again after coding arithmetic method embodies the superiority of the algorithm.

(2) GA belong to multiple search point search optimization algorithm at the same time. Many traditional research ways are using the single point research. The main drawback is that this way of searching algorithm is easy to fall into a local minimum point, thus algorithm can't out of the global minimum point. GA using multiple search point and search optimization way to overcome this shortcoming. The algorithm can carry on the global optimization in the solution space, at the same time, based on fitness function<sup>[12]</sup>, to avoid the blindness of optimization. Therefore, the GA combines advantages of directional search and random search, avoid most traditional optimization ways based on objective function or much higher derivative to defect of easy to fall into local optimum. It can obtain better area search space and the balance of the extension.

(3) GA to determine the search direction only depends on the fitness function value. Some optimization algorithm not only need to use the objective function value, and often need to objective function value of the derivative value and some other auxiliary information to determine the search direction. Genetic algorithm only need to fitness function value to determine the search direction and the way to determine the search direction can be the optimal search empty asked high positioning and the fitness function value of the solution space, to improve the search efficiency to a certain degree.

## 4.2 Genetic neural network<sup>[12]</sup> model

### 4.2.1 The genetic neural network<sup>[12]</sup> concept

Combining the BP neural network(BPNN)<sup>[1]</sup> and GA there are three commonly used ways: first is to use GA to optimize the BP neural network<sup>[1]</sup>'s connection weights<sup>[6]</sup> and threshold value<sup>[6]</sup>;The second is to use GA to optimize the BPNN's network topology;The third is to combine the first two way together.In this paper, we use the first application.It is using GA to optimize the BPNN weights<sup>[6]</sup> and threshold<sup>[6]</sup> to overcome the inherent shortcomings of BP neural network<sup>[1]</sup>.

Combining NN and genetic algorithm<sup>[11]</sup> can make full use of both advantages.It makes a new algorithm with both neural network learning<sup>[2]</sup> ability and the strong global search ability of GA.The combination of existing work have two main directions:

(1) ANN need efficient automatic design method, GA has provided a good way: BP algorithm<sup>[11]</sup> is based on the gradient method at the same time.The slow convergence speed of this method,the often plagued by local minimum point.By using GA we can get rid of this predicament, GA and NN fusion is to combine GA to the NN.The NN using GA, including the design of the coupling power, the structure and the evolution of the learning rule generation.

(2) Provides the tools for GA, NN to establish a GA based on NN support, to improve the convergence of GA.

Back propagation of BPNN learning is based on the gradient descent, and thus has the following disadvantages: network slow training speed, easy to stuck in local minimum value, and poor global search ability.Search throughout the entire solution space of GA, so easy to get the global optimal solution, and the GA does not require the objective function is continuous and different.It don't even require objective function have explicit function form, only ask questions can be calculated.

Therefore, is good at global searching capability of GA and local optimization of BP algorithm<sup>[11]</sup>, can avoid falling into local minimum value<sup>[7]</sup>, improve the rate of convergence of the algorithm, quickly get the global optimal solution of the problem.

#### 4.2.2 The combination of GA and NN

Artificial neural network(ANN)<sup>[15]</sup> and GA are applied biology to the bionics theory research results of computational intelligence. GA inspired from the nature of biological evolution mechanism, and the ANN is to the human or animal several basic characteristics of abstraction and NN simulation. As a result, they are on the way of information processing and on time there is a bigger difference. Usually, the change of the nervous system is faster, and biological evolution is need in terms of the scale of the generation. The GA and ANN have their own characteristics and strengths. In recent years, more and more scholars attempt to combine GA and ANN, hope that through the organic combination of both, take both side advantages and looking for a more effective approach to solve the problem. In general, the combination of GA and ANN can be conducted on three levels, that is the optimization of NN connection power, the NN structure optimization and the optimization of NN learning rules.

##### 4.2.2.1 GA optimization of connection weight of NN

The basic principle of the combination way is a fixed network structure, network weight can be trained using GA. Because the forward method using BP network<sup>[11]</sup> training could fall into local minimum, etc, can use the GA instead of BP as a learning algorithm of feed forward network. With GA as the learning algorithm should be mainly solve the problem of coding scheme, namely the interaction between the network weights and chromosome mapping problem.

By using GA to optimize NN connection weight<sup>[6]</sup> algorithm steps are as follows.

(1) Selected network structure and learning rules<sup>[15]</sup>. Randomly generated a set of weights<sup>[6]</sup> by using some kind of coding scheme for each code weight value<sup>[6]</sup>. It will be in the network weight value to form code chain

lined. Each code chain<sup>[9]</sup> represents a weight distribution of the network state, a set of code chain represents a set of different weights of NN;

(2) Calculated under the each corresponding chain codes of the neural network error function, which gives the genetic algorithm the fitness function, the smaller the error the higher the fitness value;

(3) Choose a number of fitness function<sup>[13]</sup> value of the largest individual form male parent;

(4) Using genetic operation like crossover and mutation operator to work on the present generation group. To produce a new generation groups;

(5) Repeat the above steps (2) ~ (4), evolution, the weight value distribution until training goals are met.

Based on GANN connection right of evolutionary computation compared with the BP algorithm<sup>[11]</sup> based on gradient descent, has the following characteristics.

(1) Genetic evolutionary method can achieve global searching, don't need error function gradient information, don't need to consider whether or not the error function different, that is the prominent advantages of it. If easy to get gradient information<sup>[17]</sup> in the training process, BP algorithm<sup>[11]</sup> may be superior to the genetic evolution method in speed.

(2) The results of the two algorithms are used in calculation process of algorithm parameters are very sensitive. The result of BP algorithm<sup>[11]</sup> of network initial state yet has a close relationship between the compliance.

(3) Genetic evolution method is good at whole searching, and the BP algorithm<sup>[11]</sup> is significantly more effective in local search.

(4) Accuracy coding limitations; genetic evolution method is sometimes hard to get the training for very high accuracy.

To sum up, if the two algorithm combined constitute a so-called hybrid training algorithm<sup>[17]</sup>, are likely to get the good training results mutual complement each other. For example, using GA to find good space region characteristic, calculated first, and initial value is given, and then using BP algorithm<sup>[11]</sup> to fine adjustment of weight, search the optimal solution. Generally speaking, the efficiency and effect of hybrid training than separate use of GA and BP training method has improved significantly.

#### 4.2.2.2 GA the optimization of NN structure

NN structure including the network topology (connection) and the node transfer function two aspects. People always expect to simple network structure to achieve the required signal processing functions, and as far as possible to achieve high performance. However, the choice of the neural network structure and design also failed to find the effective and reasonable method. At present, the most dependent on the designer's subjective experience, used for testing, comparison, to choose appropriate optimization of network structure, lack of system, the strict theoretical guidance. By using the genetic algorithm neural network can be designed according to certain performance evaluation principles, such as learning speed and generalization ability or structural complexity of the search space of optimal structure that could satisfy the requirement of the problem.

By using genetic algorithm to optimize the structure of the NN algorithm steps are as follows:

- (1) Randomly generated a number of different structure of NN, the structure of each code, each code corresponding to a network structure, N code chain constitute a population;
- (2) Using a variety of different initial connection weights<sup>[6]</sup> of each network was trained respectively;
- (3) Calculated under the each corresponding chain codes of the NN error function, the use of error function<sup>[13]</sup> or other strategies, for example network generalization ability and structural complexity) to determine each individual fitness function;
- (4) Choose a number of fitness function<sup>[13]</sup> value of the largest individual form male parent;
- (5) Using genetic operation like crossover and mutation operator to work on the current generation of group, produce a new generation of group;
- (6) Repeat the above steps (2) ~ (5), until an individual in the group (corresponds to a network structure )can meet the requirements.

#### 4.2.2.3 GA to optimize the NN learning rules

In NN training problem discussed earlier, the learning rules are set in advance, not necessarily reasonable. GA can be used to design the ANN learning rule, evolved to adapt to the requirement of environment, also can

discover new rules. Generally speaking, the evolution of the learning rule include learning parameters and learning rules of the evolution of the two aspects. To achieve this evolution, the key problem is to solve how to coding rules to learn.

By using GA to optimize NN learning rules algorithm<sup>[17]</sup> steps are as follows:

- (1) By produce a number of individuals, each individual has a learning rule, using some kind of coding scheme for each learning rule coding;
- (2) Construct a training set, among them, each element corresponds to a NN, selected to determine (or random) each of the NN structure and the initial connection, then the elements of the training set separately in each learning rules of training;
- (3) Calculate the fitness of each learning rule.
- (4) Choose a number of fitness function value of the largest individual form male parent;
- (5) Using genetic operation like crossover and mutation operator to work on the current generation of group, produce a new generation of group;
- (6) Repeat the above steps (2) ~ (5), until an individual in the group can meet the requirements.

Way of evolution in the learning rules, is a simple case study the evolution of the parameters, then, the basic learning rule (e.g., connection weight adjustment steps) has been selected, only requires the optimization parameter values of the training rules. These parameters are often have important role in the adjustment of network behavior. When learning rule itself need to adapt and optimize the network environment, will be the selection of learning rules encoding scheme is more complex.

Above discussed respectively by using GA to optimize NN power, the structure and learning rules of the combination principle of three level. In addition, the combination of GA and ANN and other various forms. For example, by using the GA for feature extraction and optimization of the NN input signal; the GA to complete the function of NN input data prepossessing. The success of feature extraction for pattern classification results is important. Parameters and characteristics of GA can be used to select data sets the scale factor, to reduce the data within the class differences, expand the difference between classes, improved ANN classification performance.

#### 4.2.3 Construction of a genetic neural network<sup>[14]</sup> model

1) Encoding. GA to optimize NN weights<sup>[6]</sup> and threshold<sup>[6]</sup> there are two main types of encoding: binary coding and real coding. Binary code problems, there are two: the first problem is the weight precision of limited length coding, coding length is short, within the scope of a certain value to meet required for solving problems in some of the combination between weights, could lead to convergence; Coding through long can lead to chromosome length is longer, will make the whole evolution process of lower efficiency. The second problem is heterogeneous, the BP neural network and one-to-many corresponding relation exists between the chromosome string, this is because the connection weights of the first connection when in a different order. Yao studies confirmed that this one-to-many mapping relationship will lead to the evolution of early maturity, the same structure of different chromosomes said appeared in the population will make the crossover operation become inefficient. And real number encoding a certain extent, overcome the shortcomings of binary encoding, in order to improve the optimization precision of the algorithm, improve the computational efficiency of the models, this paper use the real number for encoding.

(2) Define the fitness function<sup>[16]</sup>. Fitness function<sup>[16]</sup> is used to evaluate the superiority of the individual, according to the size of the fitness of individual choice, in order to make sure the individual adapt to good performance have more opportunity to reproduce, to genetic good qualities, fitness function value is the basis of a certain algorithm search direction. Fitness index model in this paper is the NN output value and the actual value of the error between the sum of squares of the countdown, the bigger the error is, the smaller fitness is. the smaller the error is, the bigger the fitness is.

Fitness function<sup>[16]</sup>:  $f = \frac{1}{SE}$ , SE as the error sum of squares

$$SE = \sum_{i=1}^n (p_i - t_i)^2 \quad (4-3)$$

Among them, p as target variables actual value, t is predicted.

(3) Population initialization and parameter selection

In the use of GA to get the parameters of BPNN, through dynamic changing the size of the population to population size on the effects of the evolution for specific experimental condition choose relatively the

population size with better effect. For BPNN, in particular within the scope of randomly generated initial population taken both the initial weights<sup>[6]</sup> and threshold<sup>[6]</sup> of real number coding chromosome.

#### (4) Genetic operations

##### 1) Selection algorithm

Step1. According to the fitness function<sup>[16]</sup> value of ordered by the individuals in the solution space, order from big to small.

Step2. Individual selection probability calculation value, the formula is:

$$p_i = r(1-q)^{n-1} \quad (4-4)$$

Among them, the q to choose the best individual probability value, n arrange position for individuals in the population, the larger the value ranking the top

$$r = q / (1 - (1-q)^m) \quad (4-5)$$

Among them, m is the size of the population.

Step3. Cycle selection, selection probability value calculation, when ascending and random sequence in [0, 1] interval, in the middle is the probability value for two consecutive times elected to the next generation.

##### 2) Cross algorithm

Randomly in a population real number coding chromosome, choose two chromosomes x, Y, the specific calculation formula is:

$$\begin{cases} X' = rX + (1-r)Y \\ Y' = (1-r)X + rY \end{cases} \quad (4-6) \quad [17]$$

In the type, r as a random number between [0, 1].

##### 3) Mutation algorithm<sup>[17]</sup>

Using non-uniform mutation algorithm<sup>[17]</sup>, formula is as follows:

$$d(X_i) = \begin{cases} (b_i - X_i)[r(1-t)]^b, \text{sign} = 0 \\ (X_i - a_i)[r(1-t)]^b, \text{sign} = 1 \end{cases} \quad (4-7) \quad X'_i = \begin{cases} X_i + d(X_i), \text{sign} = 0 \\ X_i - d(X_i), \text{sign} = 1 \end{cases} \quad (4-8)$$

On the type of  $t$  is the current generation of evolution and evolution generation, the ratio between the largest  $d$  ( $X_i$ ) is  $a$  and  $b$  chromosome  $X_i$ , around the border, the current evolution, maximum evolutionary generation and shape coefficient function is equivalent to the value of  $b$ .

#### (5) Termination conditions of evolution

When the population evolution algebra to the evolution of the set algebra terminate evolution, otherwise, turn to step (4)

#### (6) BPNN training

To end up with chromosome replication after decoding of GA to the BPNN as initial value. On the basis of data are prepared using the BPNN training. The trained BPNN to forecast is achieved. Genetic neural network<sup>[15]</sup> model of flow chart as shown in Fig 4.6.

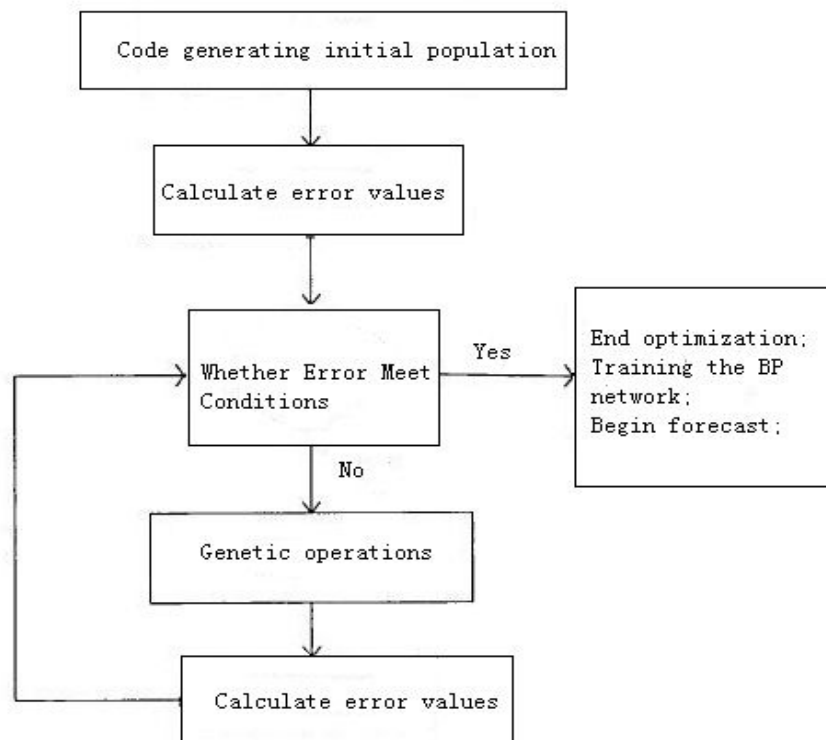


Fig 4.6 Neural Network model flowchart

### 4.3 Summary of this chapter

GA technique is applied to connection weights<sup>[6]</sup> and threshold<sup>[6]</sup> of BPNN optimization, helps to solve the BPNN in the large amount of data to predict the easy to fall into local minimum point problem. Therefore, become the key link in the short-term stock price forecast. In this chapter, on the basis of analysis of the GA, using GA to the BPNN has carried on the transformation and integration, has been based on the GA to optimize the BPNN model and the algorithm of the model and steps, lays a foundation for the establishment of the comprehensive prediction model.

## Chapter 5

### Realization and experiment of composite model of stock price forecast

In order to realize the basic function of the previously mentioned, this chapter will be the third chapter and the fourth chapter puts forward the algorithm together, constructed from the original data dimension reduction about to genetic algorithm to optimize the BP neural network prediction model of the whole experiment system model. Experimental environment with Windows XI, SP3 operating system, the experiment platform based on Matlab7.0, Matlab functions and programming language as the foundation of programming, data download software platform for wisdom.

#### 5.1.1 Overall design model

My main research focus on the principal component analysis, the GA and BPNN technology application in the short-term stock price <sup>[4]</sup> prediction. In order to achieve a certain stock closing price on the accuracy of prediction, strive to improve the BPNN based on the data to predict aspects inherent weaknesses, and to achieve a certain extent, closing price on prediction accuracy improved. At this stage for the improvement of BP algorithm <sup>[11]</sup> are mainly concentrated in two aspects: one aspect is based on the algorithm of gradient descent method was improved, such as adaptive learning rate and momentum factor method, elasticity of BP algorithm <sup>[11]</sup> etc.; On the other hand is based on numerical optimization method was improved, the commonly used a conjugate gradient method and quasi-newton method and LM method <sup>[19]</sup> etc. This chapter will be related to PCA, GA and BPNN combined with improved BPNN training, and have to be addressed in the Matlab7.0. Based on above research, this paper puts forward the comprehensive prediction model of flow chart shown in figure 5.1.

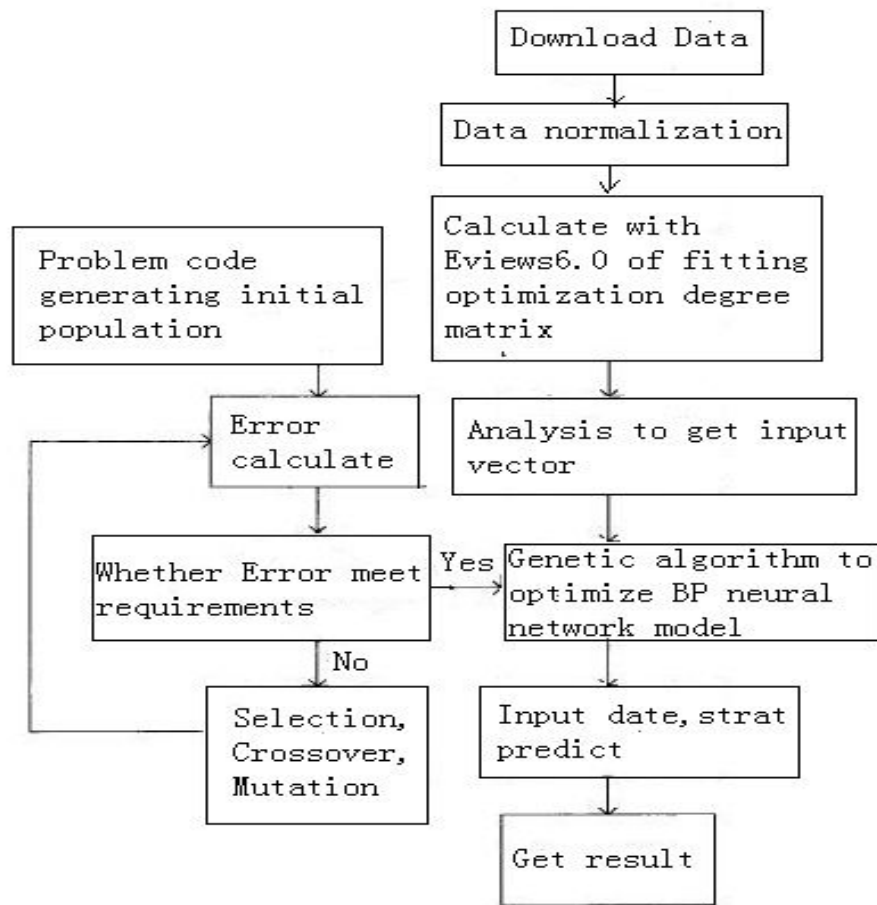


Fig5.1 Comprehensive prediction model flowchart

### 5.1.2 Summary of the design model

This chapter on the basis of the previous section completed a comprehensive prediction model is designed and realized in Matlab7.0. The model covers mentioned above two parts: a dimension reduction about module and GA to optimize the BPNN module. This part is the third chapter of this paper related to the PCA; Comprehensive prediction model for the back of the laid the foundation for empirical analysis.

## 5.2 Experiment and result analysis

BPNN technology is applied to the short-term stock price forecast factors determine exist, dimension reduction methods about selection, selection of BPNN optimization algorithm, the problems in the third chapter and fourth chapter has carried on the algorithm theory analysis and discussion, this chapter will combine the Shanghai index time series data, the article puts forward the algorithm and the establishment of comprehensive prediction model for experimental analysis respectively, through the error analysis to test the effectiveness of the proposed algorithm.

### 5.2.1 Performance evaluation index

Mean square error (MSE) <sup>[18]</sup> is commonly used in data to predict the evaluation index, and it can be said the size of error and also reflects the degree of discrete distribution error, error of smaller MSE value according to the discrete degree is smaller, the better forecast effect is. MSE computation formula is as follows:

$$MSE = \frac{1}{n} \sqrt{\sum_{t=1}^n (y_t - \hat{y}_t)^2} \quad (5-1)$$

### 5.2.2 Dimension reduction methods about contrast experiment

On the basis of the third chapter proposed algorithm, in order to compare with PCA in the aspect of BPNN dimension reduction about the effectiveness of this chapter will be two kinds of methods in Matlab simulation experiment, the Shanghai index time series data, the two dimension reduction methods about getting the input vector <sup>[6]</sup> of the input to the BPNN, for comparison and analysis from the aspects of stock price forecasting accuracy. Experiments are conducted to compare the effectiveness of the two dimension reduction methods about, the first thing to determine is the input vector of BPNN, what are the factors that affect stock price movements.

### 5.2.2.1 Experiment index selection

As mentioned in the second chapter, factors affecting the stock price<sup>[3]</sup> can be summarized as: factor, political factor, economic factor, the company industry and regional factors, psychological factors, human factors from six aspects. For stock price prediction using BPNN is the first to admit that market behavior contains all the information, all can have on stock prices volatility factors (economy, industry, psychological and artificial factor etc.) are covered in the stock price volatility. So we can use the stock price of the historical data to study the stock price changes in the future, in this, the stock price and volume reflects the basic supply and demand, is the most basic and we need to consider the core factors. Technical indicators from different angles have different classification methods, technical indicators according to the function of the technical indicators can be divided into trend indicators, overbought & oversold index, sentiment index and trend indicators four classes.

(1) Type trend indicators: the moving average (MA)<sup>[19]</sup> index refers to the recent stock price x trading day's closing price of the average.

(2) Business indicators

1) William indicators (WR). When the WR value more than 80 as oversold signal, the market has bottomed out, is a buy signal, WR value below 20 considered overbought signals, prices have peaked trend, is a sell signal.

2) Stochastic(K, D, J) is often used in the stock market short-term analysis, with the highest, the lowest price and the closing price on the basis of calculation, the main use of stock price volatility of volatility to reflect the real prices and overbought oversold phenomenon, the price has not changed before release to buy and sell signals.

3) Bias rate (BIAS). As indicators when buying and selling general setting the scope of the one is negative when the BIAS is higher than positive for a sell signal, is lower than the negative for the buy signal.

(3) Type sentiment indicators

1) Physiological line<sup>[9]</sup>, PSY values greater than 90 release a sell signal; PSY value less than 10 release buy signal.

2) On balance volume (OBV). When energy boom index and stock price change trend, this trend will continue, when there is a reverse phenomenon, the current trend reversal may occur.

#### (4) trend indicators

1) Advance decline line (ADL) <sup>[19]</sup>. ADL is the same as the share price trend, this trend will continue; If the trend, in contrast, the current trend exists the possibility of a reversal.

2) Advance decline ratio (ADR) <sup>[20]</sup>. Similar to ADL, if the stock price and trend of ADR opposite to change trend will reverse.

Each index indicators are experts in the field of financial experience and get for a long time, but also has the scope of use of each indicator and constraint conditions, with the same set of data may draw a different conclusion, therefore, in short-term stock price forecast a measure when using pure cannot insure the accuracy and comprehensive nature of problem analysis, and each index in the selection of one or more following representative index can improve the complement of the input data, can improve the quality of analysis, using the BPNN which is one of the strengths of short-term stock price forecast. In this paper, the experiment selected input vector includes representative indicators of the four categories of indicators, the indicators reflect the changes in the stock market information is relatively comprehensive, can cover all aspects of stock prediction factors. The experiment selected variables as shown in table 5.1.

Tab 5.1 The Experiment selected Variable

Variable types	Symbol	Name
dependent variable	Y	Day after Shanghai composite index closing price
independent variables	X1	Today's closing price
independent variables	X2	Yesterday's closing price
independent variables	X3	Recent 5 days average closing price
independent variables	X4	Recent 20 days average closing price
independent variables	X5	Volume of today
independent variables	X6	Volume of yesterday
independent variables	X7	Recent 5 days average volume
independent variables	X8	Recent 20 days average volume
independent variables	X9	DIF value
independent variables	X10	DEA value
independent variables	X11	K value
independent variables	X12	D value
independent variables	X13	J value
independent variables	X14	RSI value
independent variables	X15	PSY value
independent variables	X16	ADL value
independent variables	X17	ADR value
independent variables	X18	OBV value
independent variables	X19	WR value
independent variables	X20	BIAS value

Experiment of this chapter use stock price data of 60 days before 55 sets of data used for training the BP NN, the BPNN to forecast, after 5 set to predict the future 5 trading days of stock prices, with average error of the absolute value of to evaluate the effectiveness of the two groups of prediction methods, data from Shanghai wisdom network technology co., LTD., the development of the great wisdom of securities information platform.

#### 5.2.2.2 PCA of optimization for NN prediction model

As a first step, the input data matrix normalized get normalization of input data matrix as a result, the normalization methods using average method.Limited to the large amount of data, the results will be presented partial results screen shots below, on the right side of the Matlab7.0 run box using vector V1 to represent of

the normalized data matrix. After the normalization of data matrix based on the calculated correlation coefficient matrix of the original data matrix, using vector STD<sup>[14]</sup>.

The second step, calculate the correlation coefficient matrix eigenvalue and eigen-vector, characteristic value by vec, characteristic vector by val

The third step, according to the results of the calculation of characteristic root, sorting characteristic root calculation result according to the decreasing order:

14.2557, 3.5166, 0.741677, 0.523425, 0.282336, 0.196282, 0.16804, 0.15911, 0.0467774,  
0.0356118, 0.0280882, 0.0124771, 0.0113245, 0.00964113, 0.00599408, 0.00564532,  
0.000817845, 0.000290374, 0.000101947,  $8.44059 \times 10^{-11}$ .

The fourth step, calculate the proportion of characteristic root, the contribution rate, and greater than or equal to 85% as criteria for selection, experiment of principal component number four, the contribution rate of the top four, in turn,

0.71278460, 0.635614, 0.17583340281, 0.462, 0.0370, 0.838276687, 0.02617125, 0.6778381

Step 6, By principal component score matrix, in which the input vector of BP neural network, the results of detailed in the appendix as shown in Appendices 1.

(2) Identification and training of BPNN structure according to the results in the previous section, the forecast of the input matrix is 55 line four columns of the matrix, the number of principal components which are data dimension is 4, as a result, the input vector of the BP neural network for 4, output vector only after the first day of the month, the Shanghai composite index closing price, therefore, the output vector to 1.

The number of hidden layer<sup>[6]</sup> nodes is a complex problem, there is no authoritative expression to calculate the number of hidden nodes. How much hidden nodes and the number of input and output unit has a number of associations. Too little number of hidden layer nodes, not easy to make the training process of convergence, robustness can't build a strong network of predicting hidden nodes for just a few days too much, and easily makes the learning time is too long, NN training, neural network fault tolerance to lower. Currently, determine the number of hidden layer nodes use many times more NN training experience of optimal method, begin to choose the number of small training on the number of hidden layer<sup>[6]</sup> nodes of the network, and then gradually

increase the number of hidden layer<sup>[6]</sup> nodes, use the same training sample set, to determine the number of hidden layer<sup>[6]</sup> nodes network error of the hour. The existing calculation formula to calculate the number of hidden layer<sup>[6]</sup> nodes is empirical estimators, in this article, we use it as a trial and error method, the initial value of the reference. Determine the number of hidden layer<sup>[6]</sup> nodes formula basically has the following three formula:

$$\begin{aligned} m &= \sqrt{nl} \\ m &= \log_2 n \\ m &= \sqrt{n+1} + a \end{aligned} \quad (5-2)$$

In the above formula, m represents the number of hidden layer nodes, n represents the number of input layer nodes, a is constant between 1-10. The article selection formula (5.4) to determine the scope of the number of hidden layer<sup>[6]</sup> nodes is roughly, and then through trial and error method to determine the specific number of hidden layer<sup>[6]</sup> nodes.

Due to the number of input layer 4, the number of output layer is 1, therefore, the scope of the number of hidden layer<sup>[6]</sup> nodes for [3, 13]. Gradually increase the number of neurons in hidden layer<sup>[6]</sup> to the mean square error (MSE) as indicators of error, take the average of the five experiments, each experiment training 1000 times for each test, after repeated test

Test results as shown in table 5.2.

Tab 5.2 Select the training results of comparison number of neurons in the hidden layer

Hidden layer neural number	Average training error	Training time
3	0.01073170926477	1000
4	0.01216287600930	1000
5	0.00158428589832	1000
6	0.00527891931125	1000
7	0.01137761104360	1000
8	0.00420633313437	1000
9	0.08755356450830	1000
10	0.01594768396950	1000
11	0.01025209244470	1000
12	0.01141150810462	1000
13	0.02807341648376	1000

We can find that through training, under the condition of same number of training, the number of hidden layer<sup>[6]</sup> nodes when selecting 5 average minimum error, therefore, this experiment determined using BPNN structure of 4-5-1 structure.

### (3) Results

Using the momentum BP algorithm<sup>[11]</sup> of gradient descent function, the momentum factor is set to 0.9, the learning rate 0.01, the training number 1000, the BPNN, comparing the forecast value and actual value as shown in table 5.3.

Tab 5.3 Comparison of PCA of BPNN forecasting results and actual value

	1 <sup>st</sup> Day	2 <sup>nd</sup> Day	3 <sup>rd</sup> Day	4 <sup>th</sup> Day	5 <sup>th</sup> Day
Real value	2283.66	2243	2311.74	2325.68	2309.5
Predict value	1969.59	1966.87	1844.62	2026.31	2560.01

### 5.2.2.3 Related PCA to optimize the BPNN prediction model, the empirical analysis

#### (1) PCA

First of all, using the Eviews6.0 calculate the influencing factors and the second day of goodness of fit of the Shanghai composite index closing price. Calculate the goodness-of-fit of

0.944624, 0.894766, 0.20565 1, 0.894766, 0.682840, 0.667715, 0.837872, 0.888300,  
0.930419, 0.845556, 0.647644, 0.731293, 0.366749, 0.576192, 0.729064, 0.675987,  
0.540479, 0.915095, 0.544185, 0.915095.

Then according to the relevant principal component analysis steps of calculation principal component number and scoring, steps and the last section is broadly in line with PCA, to avoid repetition, here give only the final results of principal component number and see the Appendices 2.

#### (2) Identification and training of BPNN structure prediction process

Ditto the section of the experiment, the input layer number is 4, output layer is 1, therefore, the scope of the number of hidden layer nodes for [3, 13]. Gradually increase the number of neurons in hidden layer<sup>[6]</sup>, each experiment to take the average of the five experiments, training 1000 times for each test, after repeated experiments to get the training results in table 5.4.

The results from table 5.4 we still adopt 4-5-1 BPNN model for experiments.

Tab 5.4 Select training results of comparison number of neurons in the hidden layer<sup>[6]</sup>.

Hidden layer neural number	Training Average Error	Training time
3	0.00312499020211	1000
4	0.00613394702432	1000
5	0.00103293624725	1000
6	0.00518358796256	1000
7	0.00524772388272	1000
8	0.00827552030841	1000
9	0.00757423082980	1000
10	0.05896275186615	1000
11	0.06806043840070	1000
12	0.01282885874763	1000
13	0.04848087282323	1000

### (3)Results

Ditto section, using the gradient descent momentum BP algorithm<sup>[11]</sup> function, the momentum factor is set to 0.9, the learning rate 0.01, the training number 1000, to save after training the BPNN after training results and actual values, as shown in table 5.5.

Tab5.5 Related principal component analysis predicted results with actual numerical comparison of BP NN

	1 <sup>st</sup> Day	2 <sup>nd</sup> Day	3 <sup>rd</sup> Day	4 <sup>th</sup> Day	5 <sup>th</sup> Day
Real value	2283.66	2243	2311.74	2325.68	2309.5
Predict value	2024.21	1983.28	1943.46	2027.96	2064.99

#### 5.2.2.4 Error analysis

The original data, principal component analysis, the BPNN prediction results and related principal component analysis and prediction of BPNN mapping direct observation, as shown in figure 5.2, among them, the transverse coordinate unit is: the day; Unit at the vertical axis coordinates: point.

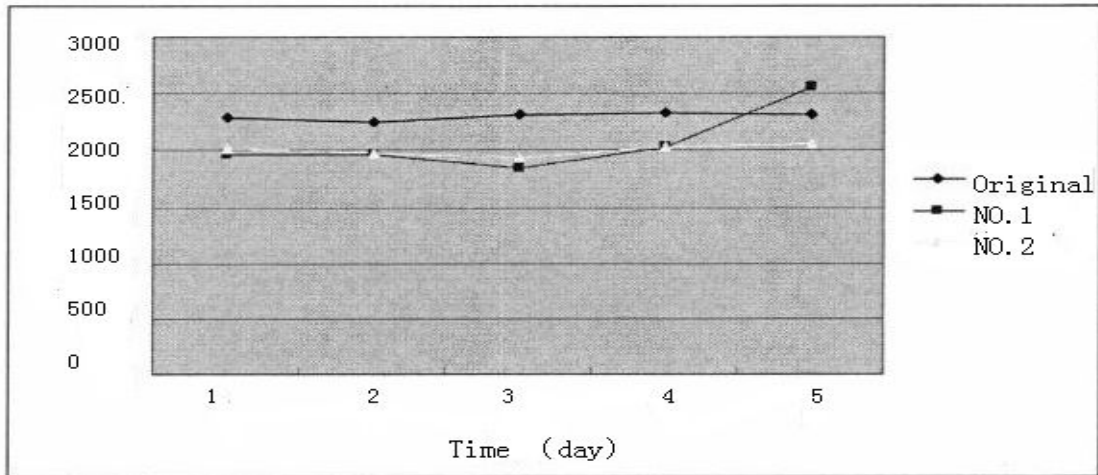


Fig 5.2 The experimental results comparison chart

From the predicted results by comparison, with principal component analysis as a dimension reduction method of BP neural network prediction about MSE\_147. 7114, with principal component analysis for dimension reduction method of BPNN prediction about MSE = 129.4286, the principal component analysis and BPNN of skeletons E value than principal component analysis and BPNN small MSE value about 12.37%, achieve the goal of the algorithm to improve prediction accuracy, suggests in this paper, PCA to improve improve the prediction accuracy of results have been achieved, PCA as a data pre-processing method of BP NN has the rationality and feasibility.

### 5.3 GA to optimize the BPNN contrast experiment

Experiments to verify in the short-term stock price forecast, genetic algorithms of BPNN optimization results, the GA is compared commonly used several kinds of optimization algorithm of BPNN prediction effect were

analyzed, in order to obtain the relevant experimental data to verify the proposed algorithm, the adopted data dimension reduction about the experimental data of experiments. In order to simplify the experiment process, more obvious comparing the effect of two kinds of prediction model, the experiment of input data in this chapter with the Shanghai composite index<sup>[19]</sup> only the opening price, the highest and the lowest price, closing price and volume as the input vector of BPNN; BPNN structure experiment with section 5.2, the determination method using 5-6-1 structure. The original data normalization after finishing see appendix 3.

### 5.3.1 GA to optimize the BPNN training process

Encoding using real number coding genetic algorithm, genetic algebra is set to 100, set the training target accuracy as 1-8, adopt LM algorithm<sup>[19]</sup> of BPNN training algorithm of experimental conditions, a research on the initial populations genetic neural network<sup>[17]</sup> prediction effect, with mean square error (MSE) as the error performance evaluation index, number of initial population were from 20 to 80 when the prediction error as shown in table 5.6.

Tab.5.6 Comparison of the prediction error of the different number of initial population table

Population number	MSE value
20	7.2926e-008
30	2.5160e-006
40	7.2794e-008
50	7.0752e-008
60	7.3122e-008
70	2.5160e-006
80	7.3026e-008

The results can be seen from table 5.6, when populations of 50 have been predicted results is best, MSE = 7.0752 e-008.

### 5.3.2 BP neural network training process<sup>[11]</sup>

#### (1) BPNN structure and training parameters

Ditto section, the BPNN is made of three layer structure. Set the training target accuracy as 1 e an 8, maximum training times, 5000 times learning rate 0.01.

#### (2) Determine the function of BP NN

BP neural network hidden layer<sup>[6]</sup> activation function use hyperbolic tangent function, training function using linear function output layer<sup>[6]</sup>.

#### (3) Determine the number of hidden layer<sup>[6]</sup> neurons

Number of hidden layer<sup>[6]</sup> neurons to choose methods and steps with the third chapter, using the momentum BP algorithm<sup>[11]</sup> of gradient descent function, the momentum factor is set to 0.9, the learning rate 0.01, the training number 1000, determine the number of hidden layer<sup>[6]</sup> is 6.

#### (4) Commonly used BPNN improved measures

##### 1) P-B connection gradient BP training function

The gradient descent algorithm:

$$x(k+1) = x(k) + ap(k) \quad (5-3)$$

$$p(k) = -g(k) + \beta(k)p(k-1) \quad (5-4)$$

Gradient P in P - B algorithm, if meet the conditions:

$$\left| g^T(k-1)g(k) \right| \geq 0.2 \|g(k)\|^2 \quad (5-5)$$

Search direction, then, would transform a negative gradient direction.

##### 2) F - P connection gradient BP training function

Compared with 1), the gradient descent algorithm of gradient transform coefficient is defined as:

$$\beta(k) = \frac{g^T(k)g(k)}{g^T(k-1)h(k-1)} \quad (5-6)$$

##### 3) P - R connection gradient BP training function

Compared with 1), the gradient descent algorithm of gradient transform coefficient is defined as:

$$\beta(k) = \frac{\Delta g^T(k-1)g(k)}{g^T(k-1)g(k-1)} \quad (5-7)$$

(4)BP training step tangent function

OSS algorithm (one step secant), the algorithm of weights and thresholds adjustment formula is:

$$w_{k+1} = w_k + c_k P_k \quad (5-8)$$

On the type of P for the direction of the search, C to reduce the search direction gradient

5) LM - BP training function<sup>[17]</sup>

LM algorithm<sup>[19]</sup> using the first-order approximate second-order Jacobian matrix Hessian matrix, the algorithm of weights and thresholds adjustment formula is:

$$w_{k+1} = w_k - [J^T J + uI]^{-1} J^T e \quad (5-9)$$

On type, J is the error of weight differential Jacobianl matrix, e is the error vector, u is a scalar.

6) Elastic BP training function<sup>[17]</sup>

Algorithm aims to eliminate the negative impact of the gradient magnitude, therefore, in carries on the correction weights, using only the partial symbol, but does not affect the rights and the amplitude value of correction, weight change depends on the size of has nothing to do with the amplitude value of revised. When the continuous phase at the same time, the iterative gradient correction value multiplied on a incremental factor, accelerates a correction; When continuous iterative gradient on the contrary, the fixed value multiplied by a reduction factor, reduce correction; Maintains a constant is zero, the specific formula is as follows:

$$\Delta x(k+1) = \begin{cases} \Delta x(k) \times k_{inc} \times \text{sign}(g(k)) \\ \Delta x(k) \times k_{dec} \times \text{sign}(g(k)) \\ \Delta x(k) \end{cases} \quad (5-10)$$

In the formula, the first k g (k) gradient iteration; K for the increment and decrements factor.

7) BFGS quasi Newton BP training function[16]

BFGS algorithm of weights and thresholds adjustment formula is:

$$M_{k+1} = M_k + \left(1 + \frac{\Delta g_k^T M_k \Delta g_k}{\Delta w_k^T \Delta g_k}\right) \frac{\Delta w_k \Delta w_k^T}{\Delta w_k^T \Delta g_k} - \frac{\Delta w_k \Delta g_k^T M_k + \Delta M_k \Delta g_k \Delta w_k^T}{\Delta w_k^T \Delta g_k} \quad (5-11)$$

On the type of M is heather approximation of matrix H, W is the network weights and threshold, g is falling gradient, k is the number of online learning.

#### 8) gradient descent momentum BP training function

The algorithm is based on gradient descent algorithm is introduced into the momentum factor, the specific formula is as follows:

$$\Delta x(k+1) = \eta \Delta x(k) + \alpha(1-\eta) \frac{\partial E(k)}{\partial x(k)} \quad (5-12)$$

$$x(k+1) = x(k) + \Delta x(k+1) \quad (5-13)$$

The ideas of the algorithm is to use the last time to influence the correction, the correction results when the last modification is too large, the modification and the last symbol, on the other hand, making the correction, reduce volatility.

#### 9) A variety of commonly used BPNN training algorithm

In the previous contents of the selected conditions, the structure and parameters in this section, the BP algorithm<sup>[11]</sup> is improved and some methods commonly used in the short-term stock price prediction research and application of the training of convergence curve as shown in the figure below (figure 5.3 to figure 5.9).

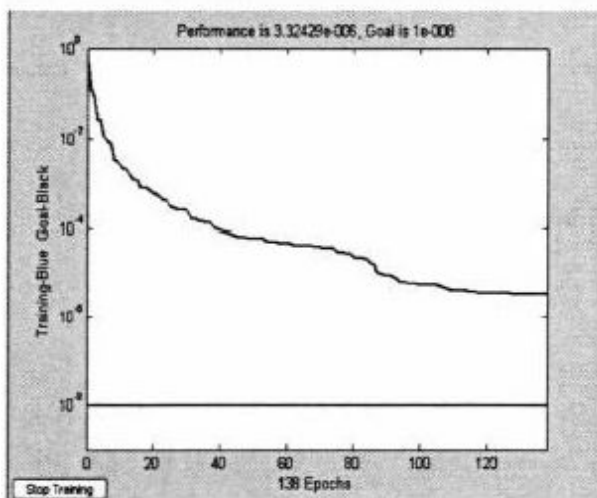


Fig5.3-B BP training function

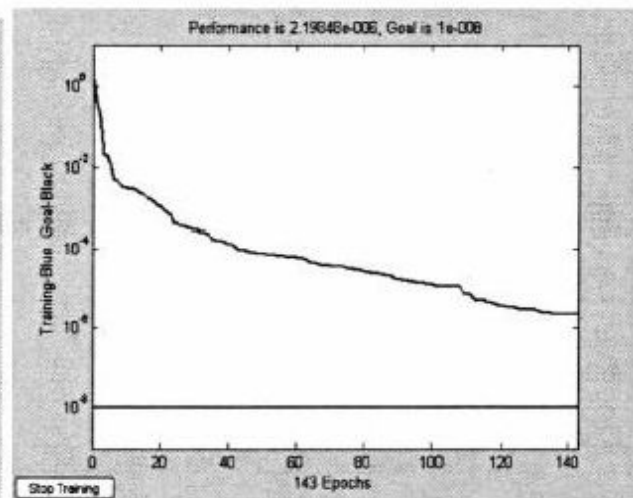


Fig5.4 F-B BP training function

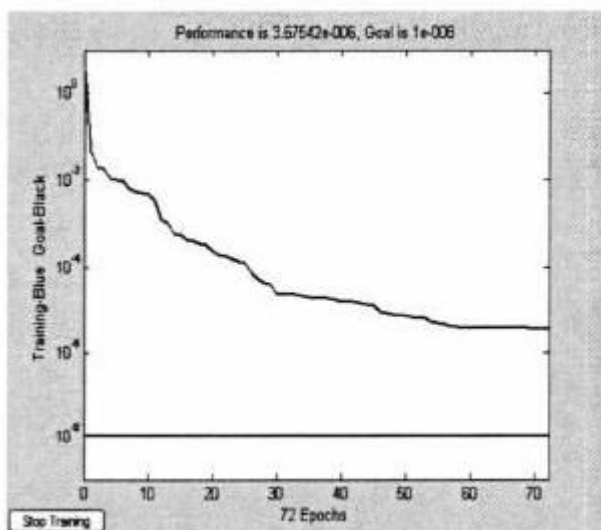


Fig 5.4 P-R Training function

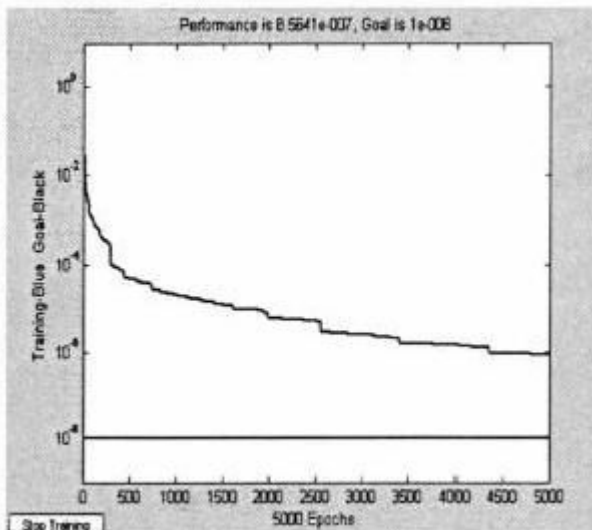


Fig 5.5 Step tangent BP training function

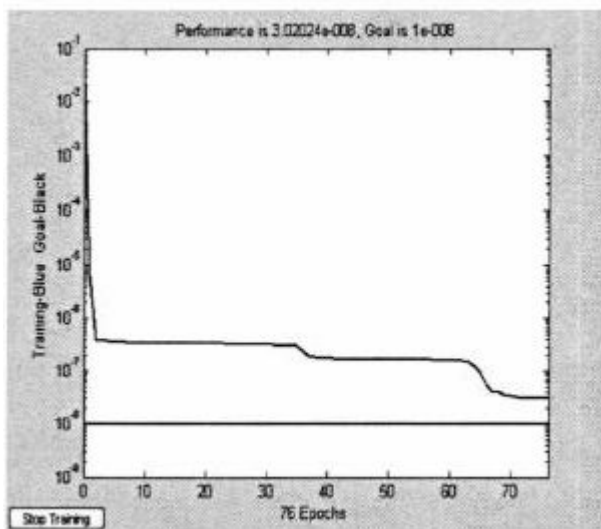


Fig 5.6 LM-BP training function

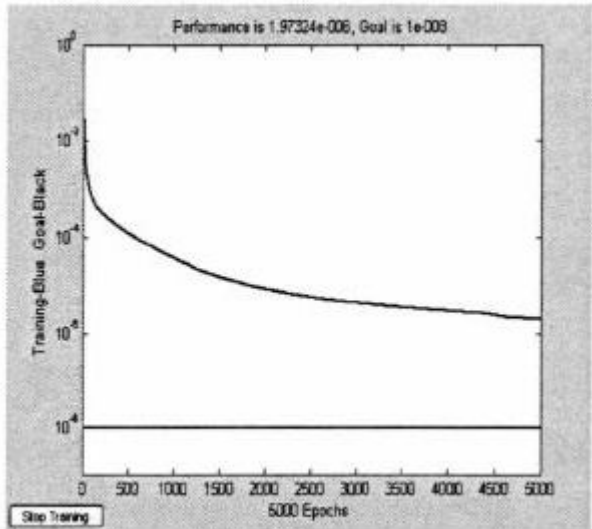


Fig 5.7 Elasticity BP training function

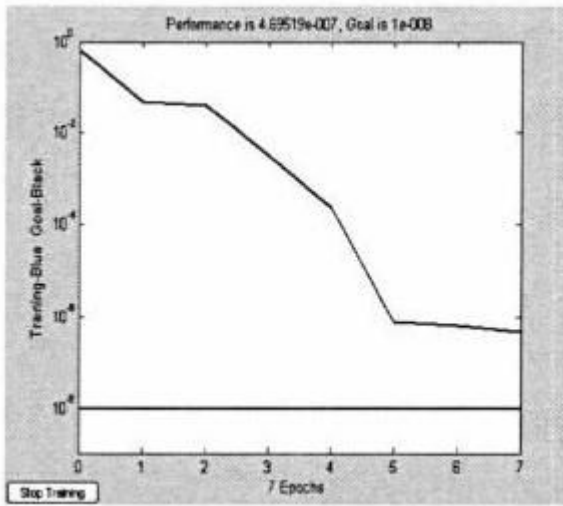


Fig 5.8 BFGS-BP training function<sup>[16]</sup>

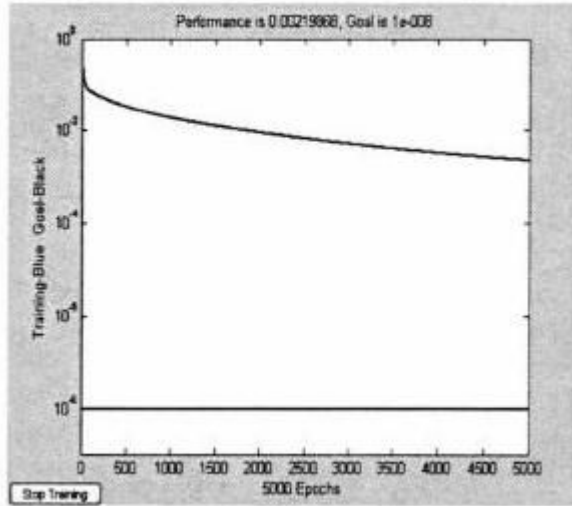


Fig 5.9 Gradient descent BP training function

Can be seen through the comparison and analysis, in view of the short-term stock price forecast, in the case of BP neural network structure is consistent, most algorithms cannot converge to target accuracy; LM algorithm in convergence speed and precision of the algorithm is commonly used in the optimal algorithm. Fully after training, the prediction results of the above several kinds of optimization algorithm error MSE arrangement as shown in table 5.7.

Tab.5.7 Kinds of improved BP algorithm forecast error comparison

BP training algorithm	MSE value
<u>Powell-Beale</u> Connect the gradient BP function	1.2341e-006
<u>Fletcher-powell</u> Connect the gradient BP function	1.7654e-007
<u>Ploak-Ribiere</u> Connect the gradient BP function	5.1860e-006
Step tangent BP function	2.7206e-007
LM-BP function	1.6521e-007
Elastic BP function	6.8166e-007
BFGS Quasi Newton BP algorithm function	4.8247e-007
Gradient descent momentum BP algorithm function	1.1079e-007

### 5.3.3 Error analysis<sup>[13]</sup>

Can be seen from the above experimental results, this group of experiment data, the genetic algebra for 50 cases of genetic algorithm to optimize the BP neural network prediction error minimum, 7.0752 e-008; And LM in traditional BP neural network optimization algorithm. The results of BP neural network is best, 1.6521

e-007. Therefore, the introduction of the genetic algorithm to optimize the BP neural network to overcome the traditional BP algorithm easy to fall into local minimum point defects. Under the condition of the same parameter Settings, the genetic neural network prediction is slightly good results with the traditional BP neural network prediction results, better fitting effect and prediction accuracy has a certain degree of improvement.

To sum up, using genetic algorithm to the BP neural network for the initial weights and threshold of optimization in the short-term stock price forecast problem is feasible and effective, the application of BP neural network in the stock price forecast provides support.

## 5.4 Comprehensive prediction model experiment

### 5.4.1 Results

In section 5.2, on the basis of experimental data were compared with the principal components analysis for the pre-treatment method of GA - BP neural network prediction model, the principal component analysis for pre-treatment method (recognition. BP neural network model and GA - BP neural network model[20] of prediction effect, the BP neural network training algorithm selects the LM algorithm, training fixed times to 1000 times. Detailed experimental steps and the former two are similar, go here, only the experimental results are given, As shown in figure 5.10, among them, the transverse coordinate unit is: the day; Unit at the vertical axis coordinates: point.

Tab.5.8 GA-BP neural network model for predicting the result of the relevant principal component analysis as a pre-treatment method

	1 <sup>st</sup> Day	2 <sup>nd</sup> Day	3 <sup>rd</sup> Day	4 <sup>th</sup> Day	5 <sup>th</sup> Day
Real value	2283.66	2243	2311.74	2325.68	2309.5
Predict value	2123.19	2083.28	2072.31	2067.58	2131.74

Tab5.9 GA-BP neural network model<sup>[18]</sup> result of the principal component analysis as a pre-treatment method

	1 <sup>st</sup> Day	2 <sup>nd</sup> Day	3 <sup>rd</sup> Day	4 <sup>th</sup> Day	5 <sup>th</sup> Day
Real value	2283.66	2243	2311.74	2325.68	2309.5
Predict value	1983.42	2011.18	1975.57	2102.23	2034.19

Tab.5.10 GA\_BP neural network model predictions

	1 <sup>st</sup> Day	2 <sup>nd</sup> Day	3 <sup>rd</sup> Day	4 <sup>th</sup> Day	5 <sup>th</sup> Day
Real value	2283.66	2243	2311.74	2325.68	2309.5
Predict value	2022.57	1993.08	2034.39	2027.96	2019.44

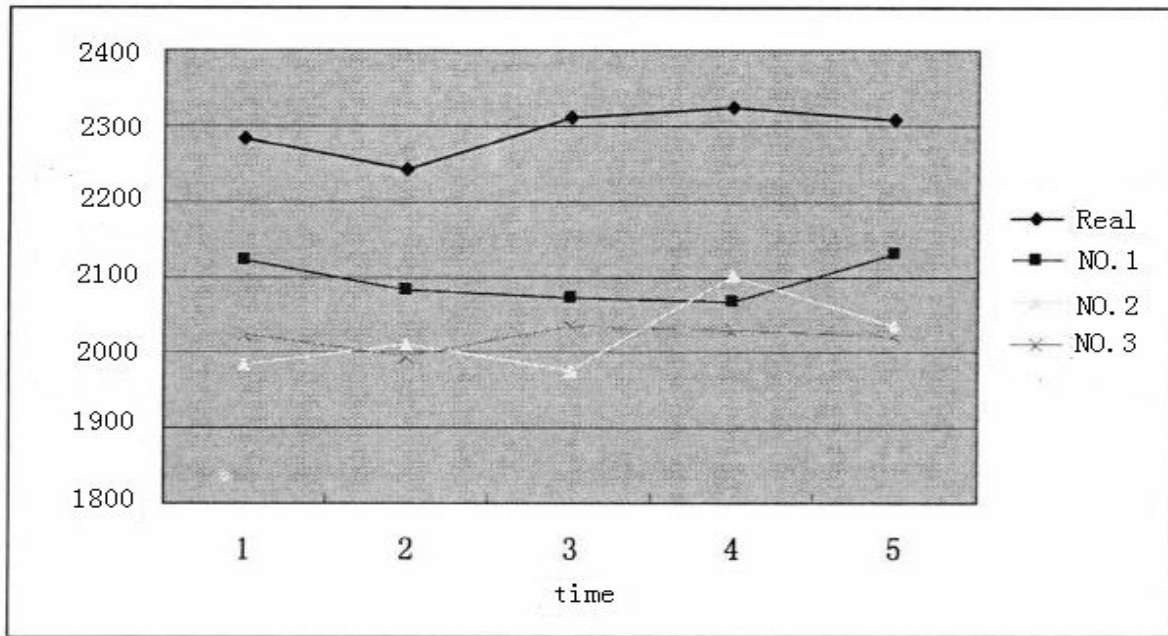


Fig 5.10 The experimental result comparison chart

#### 5.4.2 Error analysis

Three experiments of MSE<sup>[17]</sup> values as shown in table 5.11.

Table 5.11 Comparison table of forecast model error

Predict Model	MSE value
Related Principal component analysis GA-BP neural network model	90.9509
Principal component analysis GA-BP neural network model	123.7223
GA-BP neural network model	123.3406

From the result of the experiment by comparison, through the principal component analysis (pca) treated with GA - BP prediction model under the same experimental conditions, predicted results is slightly better than PCA; G- BP prediction model and GA - BP prediction model predicted results, the error precision small 26.48% and 26.26% respectively. But for a data change trend prediction and no obvious distinction, this may be due to rises to the prediction accuracy of the data is relatively easy, but with the change of stock price trends influencing factors is difficult to fully, so for the stock closing price trend change have a negative impact.

## 5.5 Summary of this chapter

On the basis of the previous section in this chapter use the Shanghai index time series data, respectively, in the previous section dimension reduction algorithm, the proposed genetic algorithm to optimize the BPNN model and the combination of comprehensive prediction model through the contrast test. On the error analysis of the experimental results of three experiments, the experimental results show that the proposed model can improve to a certain extent the existing BPNN model for short-term stock price<sup>[2]</sup> forecast accuracy, the article set up the forecasting model is effective and reasonable.

## Chapter 6

### Summary and outlook

#### 6.1 Summary

The stock market as the main body of financial market is the real economy, the role of the stock market investors have a lot of attention by government and society. Investment of the stock market all the time with risk, risk big revenue every day in the law of constant was confirmed by practice. How can you on the basis of existing a certain degree of risk control to expand revenue is very realistic. Because, the stock market inherent law of development and its regularity of prediction is an important theoretical significance and practical value. The short-term stock price forecast is a research focus in the stock market prediction is very popular in the. Stock price data is time series data, the complex in the stock market forecast system, the traditional prediction method of hard, more can't put forward and build a reasonable forecast model.

NN technology is a good deal with nonlinear and complex system model. BP neural network as one of the neural network technology, and is often used in such forecasts, and achieved some results. Therefore, on the basis of BP algorithm, this paper puts forward the improved PCA and the improved BPNN model, build a comprehensive short-term stock price forecast model, through the empirical analysis to confirm the validity of the proposed algorithm, in general, the main innovation points can be briefly summarized as the following:

- (1) According to the characteristics of the short-term stock price forecasting, the data preprocessing method to carry on the comparison analysis, principal component analysis as the core of the data processing method, and based on this joined the correlation matrix, the innovative improvements, do set the stage for the subsequent research work;
- (2) In view of the inherent in BP algorithm easy to fall into local minimum point, lack of precision, etc, combined with the genetic algorithm, the improved BPNN prediction model is given, and solved the shortcoming of BP algorithm<sup>[11]</sup> easy to fall into local minimum point, a certain extent, improve the precision of prediction.

(3) Based on the previous research has established comprehensive short-term forecasting of the stock price prediction model of the proposed algorithm and the programming implementation, using big wisdom software download the experimental data, the paper puts forward the improved algorithm with the original algorithm has carried on the empirical analysis, the result proves that the proposed relevant improvement has a certain practical effect.

## 6.2 Shortness and prospect

Stock price forecasting is a complex time series prediction problems. Is experimental and preliminary work of this paper, the work of this paper made a little improvement in terms of data accuracy, but also in terms of the change trend of stock price cannot make substantive predictions, improve prediction accuracy while at the same time, but there is a certain gap between the actual need, should be put into practice to apply still has a long way to go, still need to do a lot of specific work to increase the prediction effect of the system. Therefore, this paper argues that some problems still need to continue to explore and deepen, mainly include:

(1) For improving the prediction precision. On the one hand, this paper carefully analyses the now stock analysis in the field of technical indicators and represents the significance of a comprehensive selection of the influencing factors, and its quantitative, as the input vector of the forecasting model. However, the factors affecting the stock price fluctuation of complicated, how can a more comprehensive and effective positioning its influence factors, factors affecting how effective quantitative, will have an effect on the final prediction results. BP neural network, on the other hand, the back period of training data is difficult to determine, as the change of stock type back period also have corresponding change, but did not present a follow rule, only for a specific data was established by the method of trying, the optimal back period of training data is also in this paper, we study the determination method, points out the deficiency of the above these work also need further analysis in future research.

(2) Based on GA to optimize BP algorithm<sup>[11]</sup> parameters. Genetic algorithm and BPNN model involves many specific parameters. Such as population size, number of iterations, encoding, the number of hidden layer<sup>[6]</sup> nodes, etc. This article only by trial and error method to select the suitable parameters of certain data. These parameter Settings are as the change of data presents a certain regularity, under the condition of the

type of data is the parameter selection of the current study did not touch, and improve the model prediction effect is an important aspect, still need to do a lot of work to improve in the future.

Overall, the study of short-term stock price prediction, will be as was constantly promote and raise the economic and social development, also for the development of the stock price short-term forecasting method offers many new opportunities. And the BP neural network and its improved model will subsequently and continuous development, in the vast data of stock market, must have certain we unknown valuable information.

## Appendices

Appendices 1 BP neural network input vector principal component score matrix

0.288333148	0.012482951	0.016105067	0.00241671
0.240330926	-0.031125604	-0.000916695	-0.010902187
0.239029328	-0.03761676	-0.003503604	-0.011152829
0.198303406	-0.06979603	-0.017027208	-0.018332444
0.124944784	-0.12097954	-0.03835155	-0.033406019
0.106095758	-0.118384239	-0.03492631	-0.029469315
0.112772407	-0.093997277	-0.022546789	-0.016546284
0.132948003	-0.067166627	-0.011103969	-0.00529938
0.219994836	0.009366623	0.020121203	0.023344283
0.260160787	0.035139535	0.028636521	0.02917401
0.250481978	0.020550929	0.021721748	0.019211156
0.227378571	-0.0059496	0.009987813	0.006040745
0.216463277	-0.01728091	0.005122817	-0.000624743
0.131669009	-0.08606661	-0.022947758	-0.025150322
0.118696136	-0.0851437	-0.021882756	-0.021164278
0.146273709	-0.055489584	-0.008375431	-0.008947533
0.084963815	-0.096644702	-0.02518033	-0.022091338
0.104423494	-0.067974127	-0.01203202	-0.010081095
0.053508987	-0.092315535	-0.020779384	-0.017577296
0.031044453	-0.095934214	-0.022502099	-0.020857558
0.039526447	-0.072793801	-0.012664539	-0.012274934
0.027074253	-0.065095084	-0.009297722	-0.008669751
0.080129294	-0.013725421	0.01130109	0.00944008
0.05794791	-0.026792855	0.004815973	0.002843421
0.085268051	-0.000521686	0.015678398	0.011332161
0.063108573	-0.017057482	0.008356879	0.002649679
0.00931812	-0.058604285	-0.009508628	-0.013027118
-0.022688155	-0.075765848	-0.017194776	-0.020933789
-0.0371322	-0.073854983	-0.016046384	-0.019403236
0.00686442	-0.027566474	0.002663129	-0.001549464
-0.023630051	-0.045422554	-0.005475398	-0.007526304
0.026602625	0.001155481	0.013477473	0.009342039
0.167586368	0.108432854	0.054998341	0.04921381
0.176092447	0.092348743	0.044156595	0.035899431
0.2314637	0.120261038	0.05542506	0.046575227
0.2787231	0.125600544	0.054061359	0.047445675

0. 249796464	0. 072820057	0. 03110709	0. 022647806
0. 25446083	0. 052330081	0. 02419624	0. 011166029
0. 213638336	0. 000137972	0. 003499121	-0. 010102197
0. 363881926	0. 105984162	0. 046523329	0. 045227974
0. 402336871	0. 089164086	0. 033675259	0. 038714055
0. 397206521	0. 059266411	0. 023880432	0. 028572875
0. 383335219	0. 025648242	0. 013209676	0. 014996261
0. 388012375	0. 006196259	0. 006705741	0. 007926385
0. 341967159	-0. 05023546	-0. 015638187	-0. 012979294
0. 350942935	-0. 050759073	-0. 012785899	-0. 011126736
0. 449998006	0. 017748053	0. 015300355	0. 024300961
0. 477550097	0. 010094671	0. 008419768	0. 020564045
0. 447165666	-0. 032471343	-0. 008468588	0. 006073044
0. 490285915	-0. 008543357	0. 002596589	0. 016845039
0. 537015656	0. 010967192	0. 011244713	0. 027482791
0. 536768589	-0. 013817776	1. 14875E-06	0. 019439393
0. 547889115	-0. 023626443	-0. 003831229	0. 015248901
0. 523206537	-0. 057652255	-0. 01718522	0. 002163969
0. 51377418	-0. 077951239	-0. 025933676	-0. 006357199

# Appendices 2 Relevant principal component analysis

0.193208547	-0.01483	0.008002	0.005092
0.17785782	-0.0005	-0.0065	-0.00091
0.179576506	-0.00352	-0.00597	-0.00158
0.163966447	-0.01547	-0.00343	-0.00595
0.13268738	-0.03331	-0.00047	-0.01381
0.116439927	-0.03431	-0.0003	-0.01254
0.109287207	-0.02818	-0.00016	-0.00696
0.111310882	-0.01971	-0.00023	-0.0021
0.140452822	0.005611	-0.00103	0.01118
0.158688578	0.015476	-0.00169	0.01397
0.159008014	0.012172	-0.00237	0.010473
0.154724826	0.005014	-0.00282	0.005281
0.152160025	0.002049	-0.00326	0.002776
0.121197766	-0.02018	-0.00194	-0.00822
0.110549401	-0.02116	-0.00078	-0.00749
0.116832211	-0.01198	-0.0017	-0.0019
0.090216307	-0.02554	1.89E-05	-0.00806
0.090734426	-0.0169	-0.0006	-0.00289
0.064010083	-0.02511	0.000448	-0.0059
0.048992579	-0.02412	0.000959	-0.00602
0.044183998	-0.01622	0.000854	-0.00212

---

0.031176406	-0.01375	0.001485	-0.00076
0.046851049	0.003919	0.000484	0.00751
0.036770947	0.000875	0.000802	0.004487
0.045109941	0.010182	-0.00033	0.008615
0.036750454	0.005975	-0.00051	0.00522
0.016663607	-0.00589	0.000625	-0.00132
0.001064457	-0.00951	0.00144	-0.00396
-0.010684405	-0.00865	0.001821	-0.00333
0.000425441	0.006623	0.001634	0.004401
-0.013621648	0.001421	0.002846	0.001448
0.002009063	0.016978	0.001906	0.008882
0.057969439	0.053244	0.001523	0.026878
0.072858613	0.050639	0.001054	0.020726
0.102077112	0.058666	-5.9E-06	0.025117
0.1356915	0.060569	9.83E-05	0.023918
0.139864029	0.044593	-0.00074	0.013648
0.152998253	0.037585	-0.00365	0.009658
0.147016236	0.020223	-0.00435	0.000819
0.209151678	0.047339	-0.00169	0.020495
0.246151311	0.042235	0.000235	0.015198
0.256387695	0.029166	-0.001	0.010532
0.261525026	0.015593	-0.00306	0.005137
0.27403775	0.007635	-0.0041	0.002573
0.265646762	-0.01131	-0.00353	-0.00571
0.271153884	-0.01487	-0.00497	-0.00528
0.312327433	0.002581	-0.00336	0.007766
0.336368025	0.000293	-0.0021	0.004768
0.333605014	-0.01478	-0.00171	-0.00171
0.353876458	-0.00968	-0.00238	0.00235
0.37943096	-0.00644	-0.00275	0.005735
0.390969108	-0.01525	-0.00196	0.001811
0.403449278	-0.01959	-0.00263	-0.00034
0.400720724	-0.03117	-0.00262	-0.00551
0.402920335	-0.0373	-0.00255	-0.00877
0.40778151	-0.03833	-0.00282	-0.00863
0.373463263	-0.06134	-0.0013	-0.01828
0.409898364	-0.03332	-0.0006	-0.00092
0.429841038	-0.02865	0.00107	0.00055
0.419211138	-0.03943	0.00212	-0.00554

---

### Appendices 3 Original data normalization

0.016738	0.016757	0.016791	0.016811	0.011461	0.0166433
0.016845	0.016743	0.016791	0.016667	0.013857	0.016655421
0.016646	0.016647	0.016791	0.016679	0.013895	0.016541997
0.016708	0.016685	0.016791	0.016565	0.01628	0.016263592
0.016592	0.016476	0.016791	0.016286	0.014986	0.016206368
0.01629	0.01621	0.016791	0.016229	0.010225	0.016233209
0.016269	0.016271	0.016791	0.016256	0.011547	0.016284608
0.016284	0.016228	0.016791	0.016307	0.011348	0.01656443
0.016355	0.016534	0.016791	0.016588	0.015802	0.016663765
0.016635	0.01661	0.016791	0.016687	0.014595	0.016639994
0.016697	0.016642	0.016791	0.016663	0.013638	0.016576788
0.0167	0.016575	0.016791	0.0166	0.013601	0.016574663
0.016611	0.016558	0.016791	0.016598	0.01132	0.016305309
0.016523	0.016401	0.016791	0.016328	0.012696	0.016286104
0.016327	0.016281	0.016791	0.016309	0.010337	0.01636639
0.016348	0.016299	0.016791	0.016389	0.010808	0.016119391
0.016419	0.016288	0.016791	0.016142	0.011646	0.016178662
0.016175	0.016118	0.016791	0.016201	0.009169	0.015980858
0.016161	0.016095	0.016791	0.016003	0.009638	0.015858382
0.016011	0.0159	0.016791	0.015881	0.009452	0.015876092
0.015906	0.01582	0.016791	0.015898	0.008641	0.01581265
0.015963	0.015861	0.016791	0.015835	0.008118	0.015981094
0.015866	0.01592	0.016791	0.016004	0.010158	0.015865308
0.015969	0.015857	0.016791	0.015888	0.008689	0.015957953
0.01596	0.015949	0.016791	0.01598	0.009828	0.01587987
0.015986	0.015889	0.016791	0.015902	0.008537	0.015672936
0.015897	0.015787	0.016791	0.015695	0.010796	0.015534009
0.015679	0.015557	0.016791	0.015556	0.009629	0.01545506
0.015588	0.015523	0.016791	0.015477	0.009444	0.015585959
0.0155	0.015546	0.016791	0.015608	0.010846	0.015425779
0.015622	0.015583	0.016791	0.015447	0.011754	0.01554676
0.015459	0.01552	0.016791	0.015569	0.011892	0.015993609
0.015589	0.015994	0.016791	0.016016	0.022872	0.015972593
0.016035	0.015963	0.016791	0.015995	0.01553	0.016228801
0.016029	0.016191	0.016791	0.016252	0.021954	0.016401811
0.016333	0.016358	0.016791	0.016425	0.021997	0.016330419
0.016414	0.01637	0.016791	0.016353	0.0178	0.016393625
0.01638	0.016349	0.016791	0.016417	0.014977	0.016226361
0.016413	0.016313	0.016791	0.016249	0.013778	0.01692808
0.016297	0.016871	0.016791	0.016952	0.030669	0.01700451
0.017003	0.017005	0.016791	0.017028	0.027499	0.017021197

0.017033	0.017115	0.016791	0.017045	0.02404	0.017019465
0.017055	0.01705	0.016791	0.017043	0.017995	0.017067558
0.017036	0.017046	0.016791	0.017092	0.018947	0.016949175
0.017141	0.017168	0.016791	0.016973	0.019707	0.016994356
0.016992	0.017011	0.016791	0.017018	0.01496	0.01742381
0.017021	0.017402	0.016791	0.017448	0.027324	0.017467259
0.017473	0.017404	0.016791	0.017492	0.023231	0.017363123
0.017569	0.017515	0.016791	0.017388	0.024597	0.0175784
0.017444	0.017516	0.016791	0.017603	0.0221	0.01786082
0.01767	0.017788	0.016791	0.017886	0.024408	0.017922688
0.018089	0.017996	0.016791	0.017948	0.026589	0.01798857
0.017948	0.017975	0.016791	0.018014	0.022527	0.017915446
0.018051	0.017942	0.016791	0.017941	0.023431	0.0179097
0.017945	0.017901	0.016791	0.017935	0.023261	0.017975189

## Appendices 4 Cwfac module

```
function result=cwfac(vector);
fprintf('correlation coefficient matrix:\n')
std=CORRCOEf(vector)
fprintf('feature vectors(vec)and eigenvalue(val);\n')
[vec,val]=eig(std)
newval=diag(val);
[y,i]=sort(newval);
fprintf('Characteristic root sorting:\n')
for z=1:length(y)
    newy(z)=y(length(y)+1-z);
end
fprintf('%g\n',newy)
rate=y/sum(y);
fprintf('\n Contribution:\n')
newratte=newy/sum(newy)
sumrate=0;
newi=[];
for k=length(y):-1:1
    sumrate=sumrate+rate(k);
    newi(length(y)+1-k)=i(k);
    if sumrate>0.85 break;
end
end
fprintf('principal component number:%g\n\n',length(newi));
fprintf('principal component load:\n')
for p=1:length(newi)
```

```
for q=1:length(y)
    result(q,p)=sqrt(newval(newval(newi(p)))*vec(q,newi(p)));
end
end
disp(result)
```

## Appendices5    Cwprint module

```
function print=cwprint(filename,a,b);  
fid=fopen(filename)  
vector=fscanf(fid,'%',[a b])  
fprintf('stander result:\n')  
v1=cwstd1(vector)  
result=cwfac1(v1);  
cwscore1(v1,result);
```

## Appendices 6      Gabpnet   program

```

function net=gabpnetl(p,t)

pa=load('D:\report\data\BPpvector.txt');
ta=load('D:\report\data\BPt.txt');
pa1=load('D:\report\data\BPpl.txt');
tal=load('D:\report\data\BPtl.txt');

p=pa
t=ta;

p1=pa1;
t1=tal;

nntwarn off

net=newff(minmax(p),[6,1],{'tansig','purelin'},'trainglm');

P=p;
T=t;

R=size(P,1);
S2=size(T,1);

S1=6;

S=R*S1+S1*S2+S1+S2;

aa=ones(S,1)*[-1,1];

popu=50;

initPpp=initializega(popu,aa,'gabpEvall');

gen=100;

[x,endPop,bPop,traceInfo]=ga(aa,'gabpEvall',[],initPpp,[1e-8
1],'maxGenTerm',gen,'normGeomSelect',[0.09],['arithXover'],[2],'nonunifMutation',[2
gen 3]);

[W1,B1,W2,B2,P,T,A1,A2,SE,val]=gadecod1(x);

net.IW{1,1}=W1;

```

```

net.Lw{2,1}=W2;
net.b{1}=B1;
net.b{2}=B2;
p=P;
t=T;

net.trainParam.show=500;
net.trainParam.lr=0.01;
net.trainParam.mc=0.9;
net.trainParam.epochs=5000;
net.trainParam.goal=1e-8;

[net,tr]=train(net,p,t);
A=sim(net,p1)
Et1-A
MSE=mse(E)

```

## Appendices7 gabpEval Program

```
function[sil,val]=gabpEval1(sol,options)

pa=load('D:\report\data\BPpvector.txt');
ta=load('D:\report\data\BPt.txt');
pa1=load('D:\report\data\BPpl.txt');
tal=load('D:\report\data\BPtl.txt');

p=pa
t=ta;

p1=pa1;
t1=tal;

nntwarn off

R=size(p,1) ;

S2=size(t,1);

S1=6

S=R*S1+S1*S2+S1+S2;

for i=1:S,

    x(i)=sol(i);

end;

[W1,B1,W2,B2,P,T,A1,A2,SE,val]=gadecod1(x);
```

## Reference:

- [1]H.white,“Economic prediction using neural networks,”the case of IBM daily stock returns.Neural Networks,IEEE International Conference on 1988.2(6),PP451-458.
- [2]T.Kimoto,K.Asakawa,M.Yoda,“Stock Market Prediction System with Modular Neural Networks,”Neural Networks,1990, International Joint Conference on IJCNN.1990.1,PP1-6.
- [3]Dogac Senol,“Prediction of stock price direction by artificial neural network approach,”research paper,2008.
- [4]Ajith Abraham, Ninan Sajith Philip and P. Saratchandran,“Modeling Chaotic Behavior of Stock Indices Using Intelligent Paradigms,”International Journal of Neural,Parallel&Scientific Computations,USA,Volume11,Issue(1&2),2003,pp 143-160.
- [5]Bishop C. M., “Neural Networks for Pattern Recognition,” Oxford: Clarendon Press, 1995.
- [6]Vapnik V., “The Nature of Statistical Learning Theory,”Springer-Verlag, New York, 1995.
- [7]Philip N.S. and Joseph K.B., “Boosting the Differences: A Fast Bayesian classifier neural network,” Intelligent Data Analysis, IOS press, Netherlands, Volume 4,2000, pp 463-473.
- [8]Jang J. S. R., Sun C. T. and Mizutani E., “Neuro-Fuzzy and Soft Computing: A Computational Approach to Learning and Machine Intelligence,” Prentice Hall Inc, USA, 1997.
- [9]G.peter Zhang,“Times series forecasting using a hybrid ARIMA and neural network model[J],”Neurccomputing,2003(50),pp159-175.
- [10]Chi-Jie LU,“Integrating independent component analysis-based denosing scheme with neural network for stock price prediction,”2010.37(3),pp56-64.
- [11]A.Murat.Bahadir,“Comparison of bayesian estimation and neural network model in stock market trading,”Intelligent Engineering Systems through Artificial Neural Networks, 2010.02,pp74~81 .
- [12]Hammda,M.alhajali,“Forecasting the jordanian stock prices using artificial neural network,” Intelligent Engineering Systems through Artificial Neural Networks,2007(17),pp273~275.
- [13]A.Murat ozbayoglu,“Neural based technical analysis in stock market forecasting,”

Intelligent Engineering Systems through Artificial Neural Networks,2008(18),pp261-265.

[14]Arnold F.shapiro,“The merging of neural networks,fuzzy logic and genetic algorithms[J],” Mathematics and economics,2002(32),pp115-131.

[15]Melike Bildirici,“Improving forecasts of GARCH family models with the artificial neural networks:An application to the daily returns in Istanbul stock Exchange[J],” Expert Systems with Applications,2009(36),pp7355-7362.

[16]Tsung-Jung Hsieh,Hsiao-fen Hsiao,Wei-Chang Yeh,“Forecasting stock markets using wavelet transforms and recurrent neural networks:An integrated system based on artificial bee colony algorithm,” Applied Software Computing,2011(11),pp510-525.

[17]Tung-Kuan Liu,Chiu-Hung Chen,Jyh-Horng Chou,“Method of Inequalities-based Multi-objective Genetic Algorithm for Optimizing a Cart-double-pendulum[J],”Automation and computing,2009.2,pp72-73.

[18]Park Y R,Murray TJ and Chen C,“Predicting sunspots using a layered perception neural network,” IEEE Trans,Neural Networks,1996.7(2),pp501-505.

[19]Whiteley D,Hanson T,“Optimizing Neural Networks Using Faster,More Accurate Genetic Search[C],” Proc.of 3<sup>rd</sup> Conf.On,GA.Arlington,1989.

[20]Holland J H,“Adaptation in Nature and Artificial System[J],”MIT Press,1975.1,pp3.