

Prediction of Share Price by Neural Network

Hassan saleh nejad

Department of Management, Economics and Accounting, Payame Noor University, I.R.of Iran

Abstract: Recently, determining how to recognize the volatility of price changes is one of the major issues that are interesting for investors and financial analysts and it led to different approaches and various solutions. Therefore, 19 financial ratios which were effective on stock price were used and Neural Network was used for predicting stock price of 393 companies listed in Tehran Stock Exchange during 1988-2010. In current study, we used SPSS, MATLAB and WEKA for testing hypotheses. Consequently, results indicate that Neural Network can predict stock price with accuracy of %91.68. In fact, significant relation exists between ability of Neural Network and predicting of stock price.

[saleh nejad H. **Prediction of Share Price by Neural Network..** *J Am Sci* 2012;8(12):1427-1431]. (ISSN: 1545-1003). <http://www.jofamericanscience.org>. 192

Key words: Neural Network, Stock price, financial ratios, Support Vector Machine, Genetic Algorithm, particle swarm optimization

Introduction:

Investment and capital accumulation have vital role in development economy of countries. Importance of this factor and essential role of investment can be clearly observed in capitalist countries. Definitely, Stock market is one of the most suitable places for absorption of small capital and using in grow of a company, macro level and grow of investor. Since, the definition of investment is postponed of consumption in order to consume more and better in future; investors expect to reach expected return. Therefore, the most important issue in this context is purchasing a share of the low price and selling it at higher prices and predicts the stock price is vital matter. After the opening of the stock markets always there have been ideas of predicting share price by software, hardware and various financial analyses. Experts of stock markets have studied the markets for sequential years and they have found patterns and predict based on these patterns. There has been various software which help for making decisions and used as predicting engineers, math methods like linear regression, moving average and so on.

Importance of share price predicting:

Lack of knowledge about changes of share price can lead to use of predicting share price of companies. Efficiency of predicting shows amount of success in predicting behavioral patterns. In fact, process of generating stock price can be studied as a model of reliability. The process may be obtained by linear models, nonlinear models and stochastic models. Now, one of attractive and important fields of investment and financial analyses is determining how to recognize the volatility of price changes is led to various solutions

editor@americanscience.org

and approaches. In this context, as result of the lack of access to accurate information and evidence about the determinants of volatility and changes of stock price predict of trend and stock price volatility is not possible easily and this problem provides efficient hypothesis; in other words, by using available information predict of share price is not possible.

The opposite point of this theory is the ability to predict stock prices and this perspective believe time series of prices is not random but follows certain behavioral patterns. In middle of 70s and particularly since 80s there have been wide efforts in order to predict share price by new math methods, time's series and non-linear methods and it can be cited test on price and stock indexes in USA and UK and all of them have tried to discover structure of time's series. Basically prediction is a key element for management decisions. In general we can say that prediction is estimate of future events and aim of predicting decline of risk in making decision. Forecasts are often incorrect and errors can be reduced errors a bit more information about the system. On the other hand, higher predicting cost can reduce uncertainty risk.

Neural Network:

Network formed of numerous of simple processors (neurons) that are widespread interconnected, operate in parallel, and learn from experience. These are the initially known characteristics of biological neural systems that are the easiest to explain in artificial neural system (Specht, 1991). Neural Network is model which process information like human brain. This model is not programmed for particular action, however with more interaction and over time becomes more experienced

<http://www.jofamericanscience.org>

and able to logical conclusion. Recent developments in the application of artificial intelligence tools, particularly in the field of artificial neural networks in financial affair have opened new approaches. The artificial neural network models derived from the data and use widely for solving various problems particularly for predicting. In according to share price and predicting, the current research investigates prediction accuracy of the neural network technique; provide information for predicting share price.

Financial ratio:

Financial information and financial ratios are considered to assess the financial position, financial ratios as predicting variables are useful for studies of corporate bankruptcy, seems Paul Fitzpatrick (1932) is the first person who detailed study to predict the bankruptcy of companies by analyzing financial ratio and he concluded that financial ratios predict bankruptcy somewhat. Beaver (1966) choose 30 financial ratios as the most indicator financial health of company and concluded that these ratios are difference in healthy and failure companies. Moreover, he found higher ability of prediction of financial distress by certain financial ratios like cash flow to total debt. Hossary (2006) by investigating 28 researches in field of predicting bankruptcy during 1966 and 2004 and he concluded that over 79% of these researches used financial ratios as predicting bankruptcy. Chen (2011) showed that financial ratios had higher performance than non-financial ratios and economical indicators.

Literature Review:

Refenes and *et al* (1994) by modeling the behavior of stock prices by neural networks, its performance has been compared with the regression models. In this by neural network is used as a replace of classical statistical analyses for predicting share prices of big companies. The result shows that neural network compare with statistical techniques have better performance and models. Tan and *et al* (1995) designed a system which predicts significant changes in short period. First, Pre-processing was performed on the data and then neural network makes a model which predicts very profitable position. Richard Oslon (1996) in his research "investigates relationship between cash and accrual components with stock future price". According to the 1980 Statement of Financial Accounting Standard Setting Board, which implies cash originated from operational activity is an instrument for controlling earnings quality. In main hypothesis indicates that stability of current earnings is more than accrual earnings. Moreover, in other hypothesis also indicate that stock price can not reflect of stability of current earnings in cash components.

editor@americanscience.org

Result of this studies shows that performance of earnings of accrual earnings has lower stability compare with cash earnings. In addition, in companies which have higher (lower) components of accrual earnings, they will have positive (negative) returns of uncommon shares.

Hypothesis:

This study sought to answer the following questions

H₁: Does neural network predict share price of companies listed in Tehran Stock Exchange?

Data and Methodology:

This study is inductive and it makes use of past information and historical financial statement. This study is also a correlative study since it seeks to investigate the relation between dependent and independent factors. It is a periodic study because it studies a specific period of time and it can be an applied research. It is a periodic study because it studies a specific period of time and it can be an applied research. Independent and dependent variables and primary processing of data were carried out by Excel. The assumption of the research is tested based on the concluded that the effect of higher-quality accounting on regression analysis with the aid of SAS and SPSS statistical analysis software. In order to gather theoretical information, library research was selected and the books in the libraries, together with articles found in the internet, were used. An empirical research was used to describe the events in Tehran stock exchange (TSE) and investigate the correlation of variable by regression analysis. The TSE listed companies were chosen as a population and then some samples were selected based on the following conditions:

- 1) The entities should be listed before 1987.
- 2) Date financial firms should lead to the end of March each year.
- 3) The entities should be activated during 1988 to 2010.
- 4) The entities should not change their financial periods.
- 5) The entities' availability of information is required.

Table 1. Independent variable

X₁	Earnings Quality
X₂	Current ratio
X₃	Ratio of Capital Working
X₄	Ratio of net Profit to Total assets
X₅	Ratio of net profit to equity

<http://www.jofamericanscience.org>

X ₆	Interest payment
X ₇	Sales growth rate
X ₈	Rate of Profit
X ₉	Cash flow on total assets
X ₁₀	Ratio of retained earnings to total assets
X ₁₁	Ratio of sales to total assets
X ₁₂	Ratio of Liability
X ₁₃	Ratio of cash flow on liability
X ₁₄	Ratio of capital working on total liabilities
X ₁₅	Ratio of cash flow on net profit
X ₁₆	Debt to Equity Ratio
X ₁₇	The ratio of operating income to total assets
X ₁₈	Equity to total assets
X ₁₉	Current debt to operating cash flow

In this equation the aim of optimization is to find values for $x(1), \dots, x(20)$ so that the equation has its minimum value. But in this research we will find these variables by neural network.

Neural Network

The term neural network was traditionally used to refer to a network or circuit of biological neurons. The modern usage of the term often refers to artificial neural networks, which are composed of artificial neurons or nodes. Thus the term has two distinct usages:

1. Biological neural networks are made up of real biological neurons that are connected or functionally related in a nervous system.
2. Artificial neural networks are composed of interconnecting artificial

An Artificial Neural Network, often just called a neural network, is a mathematical model inspired by biological neural networks. A neural network consists of an interconnected group of artificial neurons, and it processes information using a connectionist approach to computation. In most cases a neural network is an adaptive system that changes its structure during a learning phase. Neural networks are used to model complex relationships between inputs and outputs or to find patterns in data.

There is no single formal definition of what an artificial neural network is. Generally, it involves a network of simple processing elements that exhibit complex global behavior determined by the connections between the processing elements and element parameters. Artificial neural networks are used with algorithms designed to alter the strength of the connections in the network to produce a desired signal flow.

The word *network* in the term 'artificial neural network' refers to the inter-connections between the

editor@americanscience.org

neurons in the different layers of each system. An example system has three layers. The first layer has input neurons, which send data via synapses to the second layer of neurons, and then via more synapses to the third layer of output neurons. More complex systems will have more layers of neurons with some having increased layers of input neurons and output neurons. The synapses store parameters called "weights" that manipulate the data in the calculations. An ANN is typically defined by three types of parameters:

1. The interconnection pattern between different layers of neurons
2. The learning process for updating the weights of the interconnections
3. The activation function that converts a neuron's weighted input to its output activation.

When performing classification analysis with a set of existing data, one common approach, called holdout validation, is to split the data into a larger data set (often 80 percent) for training the neural network and a smaller data set (20 percent) for testing the model. Training can be defined as finding the neural network weights and biases that minimize some error value. Testing means evaluating the neural network with the best weights found during training, using some measure of accuracy.

One common approach used to test learned regularities is to divide the data set into two parts and use one part for training and another part for validating the discovered patterns. This process is repeated several times and if results are similar to each other than a discovered regularity can be called reliable for data.

Support vector machines:

Vapnik proposed the support vector machines (SVMs) (Vapnik, 1995). The support vector machines (SVMs) have been highly concerned in recent years. Based on the structured risk minimization (SRM) principle, SVMs seek to minimize an upper bound of the generalization error instead of the empirical error as in other neural networks. Additionally, the SVMs models generate the regress function by applying a set of high dimensional linear functions. The SVM regression function is formulated as follows:

$$y = w\phi(x) + b,$$

The training algorithms of SVMs try to find the optimal separating hyperplane by maximizing the margin between the hyperplane and the data and thus minimizing the upper bound of the generalization error. Delivering promising results makes the SVMs extensively applicable in many information processing

<http://www.jofamericanscience.org>

According to the table, numbers from -10 to 10 by step 1 are given to each variable. Therefore total conditions that should be investigated by MATLAB are 20^{19} . One of these conditions (or more if their values are equal) is the ideal number that if we multiply each column of excel file by this number, neural network will perform the best prediction.

After multiplying and creating new data matrix, necessary raw data is available for neural network. Therefore, we select 75% of the rows randomly and use them for training, and use the remaining 25% for prediction, and repeat this action 10 times. These numbers show percent of accuracy of obtained answers through the processes. By calculating the mean of these 10 numbers, MATLAB find the correct answer. This number is saved in a 19 dimensional matrix such as $A(1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,1,6)$.

Now by using several for loops, we find both the 19 variables that provide us the most accurate estimation and the percent of correct answer prediction.

Conclusion:

Based all of that has been said, if we have less ascending trend, we will have more accuracy. However, Due to the long calculation, we used the first trend in this study and final result is equal to 2.66. It means the absolute difference between the predicted responses with the correct answer is 2.66 and in the most ideal situation should be zero. Furthermore, due to all data in real time and predicted time are the same. Definitely, it can be said the average distance between the actual and predicted data is 2.66 and this number is negligible compared to the original value that definitely change in the trend, we will have more accuracy. Finally, the entire study was repeated and there was not any significant change in result of the study. It means remove of low value columns may accelerate problem solving. Eventually, the results indicate the existence of errors in forecasting stock prices was 8.32 percent. In other words, the model able predict stock price by 91.68 percent.

Future research:

Due to the high accuracy prediction Neural Network is proposed that PSO or SVM combine with other models like honey bee colonies and also it is

better to use of market price and compare it with the results of this research. Moreover, it is proposed that investigate, is there any difference between variable methods lead to the selection of the different independent variables? And also check whether or not will change the number of these variables?

Reference:

1. Beaver, W., (1966), 'Financial ratio as predictors of failure, empirical research in accounting selected studied 1966, Journal of Accounting Research,4,pp71-111
2. Chen, M., Y. (2011), "Bankruptcy prediction in firms with statistical and intelligent techniques and a comparison of evolutionary computation approaches", Computers and Mathematics with Applications, 62 ,PP: 4514–4524
3. Kennedy, J. Eberhart, R, (1995),"Particle swarm optimization", IEE: International conference on Neural Network, pp: 1942-1948.
4. Fitzpatrick, p., (1932)," A comparison of ratios of successful industrial enterprises with chose of failed companies", The certified public Accountant, October, November,Decembery, pp:598-605, 656-662, 727-731,respectively.
5. Ohlson, J., (1980)," Financial ratios and the probabilistic prediction of bankruptcy", Journal of Accounting Research, 18(1), pp109-131
6. Hossari,G.,(2006),"A Ratio-Based Multi-Level Modeling Approach For Signaling corporate collapse: A study of Australian corporations" , Swinburne university of Technology ,Australia, A thesis for the degree of Doctor of philosophy.
7. Specht, D (1991), A General Regression Neural Network, IEEE Transactions on Neural Network, Vol. 2(6):568-576
8. Trelea, I, C (2003) the particle swarm optimization algorithm: convergence analysis and parameter selection, Information Processing Letters 85" 317–325
9. Vapnik V (1995), the nature of statistic learning theory. New York: Springer
10. Wu, K, P., Wang, S, D (2012), Choosing the kernel parameters for support vector machines by the inter-cluster distance in the feature space, Pattern Recognition, Vol. (42): 710 – 717.

11/21/2012