International Journal of Artificial Intelligence and Neural Networks – IJAINN
Volume 4 : Issue 1 *[ISSN 2250 – 3749]*
Publication Date : 09 January 2014

Using Back Propagation Neural Network Techniques on Prediction of Taiwan Stock Market

Jin-Cherng Lin, Fu-Jyh Ko, Jian-Hon Liu, and Yueh-Ting Lin

ABSTRACT -- In this research, we use Back Propagation Network method combined statistics to measure up the analysis on the close value of stock market, and individual stocks and predict the trend of stock and individual stocks index in a period of time, also use RMSE and daily average error to analyze result. During this research, we use the close value of Taiwan Stocks Index from 1999 to 2010, and for the individual stock, we choose Taiwan Semiconductor Manufacturing Company (TSMC) and ASUS Company to perform the test. Based on the training data of the Back Propagation Network and prediction data during this experiment result, we find out use Back Propagation Network method combined statistics can achieve certain degree result on assessment.

Keyword -- Stock-index estimation, Back propagation neural network, Back propagation network algorithm, Root-mean-square-error.

1. Introduction

In the application fields of computing science,there has been long-lasting progress in techniques and its applying of the computing intelligence. It is often used to make forecast of various statistic data. Furthermore, according to the researches, it is very effective to make unpredictable forecasts for the diverse data threads.

Among enormous applications, the projection of the stock market is regarded as a very challenging task. In tradition, all the projections of the stock share index and fluctuation are analyzed by human experts. The specialists use their personal experiences to estimate the fluctuation of index number and price of the coming days.

Basically, the time series of financial data those influence the stock market is very dynamic and tend to be affected by news. Besides, it is not linear and combines many external factors. There might be some difficulties using traditional statistic regression analysis to solve these problems. However, with the using of computing intelligence techniques, such as neural network sensitivity analysis [Zurada 1994, Hashem 1992, Ghosh 2008, Gevrey 2003, Gevrey 2006] and neuro-fuzzy approaches [Chen 2011, Chen 2010, Wong 2010, Chen 2009, Chen 1996, Chen 2002, Chen 2006], better results of analyzing and prediction for the stock market can be expected [Wu 2012, Armano 2005, Atsalakis 2009, Hadavandi 2010, Shao 2012, Tsang 2009, Kim 2006].

In this article, a study on combining statistics and Back Propagation Network (BPN) algorithm [Rumelhart 1986] to estimate the probable trend of stock market in Taiwan is proposed. In addition, owing to the real target of our business in the stock market being the specific share, it is more practically valuable to the projection on the price of specific share. We thus apply the estimation to the specific share too. The accuracy of estimation is also validated strictly.

Jin-Cherng Lin et al.
Dept. of Computer Science & Engineering, Tatung University,
Taipei 10452, Taiwan
jclin@ttu.edu.tw

The BPN is the major technique to produce prediction data in our approach. The historical data will be pre-processed to fit the format of BPN algorithm. In additional to the data of main factors, (i.e., weighted index of the Taiwan stock market), the data of secondary factors (i.e.,Dow Jones index, Nasdaq index, and M1B number in Taiwan) are gathered to be the input of BPN algorithm in chronological order. It is done to pursue high accuracy of the prediction because the observations show the Taiwan stock market is affected strictly by these secondary factors.

The weighted index of Taiwan stock market and prices of specific shares in market closing daily are focused to make forecast. The outcome of this research shows that the accuracy of forecast is over 60% in normal situation of financial market. However, the artificially selection is still needed to obtain higher accuracy in some special occasions such as financial tsunami and natural disasters.

2. Related Works

F.Rosenblatt(1957) claimed that sensor structure is the most ancient neural network, but this pattern is the single layer network structure which is in lack of the learning algorithm of the hidden layer, resulting in inability to solve Exclusive OR and limited learning algorithm. While Dr.P.Werbos reported the learning algorithm of the hidden layer called back-propagation network, BPN. Not until D.Parker again proposed BPN in 1957 and D.Rumelhart, G.E.Hin-ton, and R.J.Williams reported an article on it did it became widely spread [Rumelhart 1986].

The research focuses on 1O shares of MSCI, using technical and basic analysis, including smooth comparison movement average line, KD,J random index line, RSI relative intensity index, BIAS deviation rate, William index WMS%R, moving average line, Psychological line index, momentum index, tendency index, and stop loss point index along with simultaneous information to be the variable of the research. The data will be led into the model which consists of back propagation network to produce estimate value to test and verify it [Huang 2009].

The approach in this research is based on fuzzy time series and fuzzy variation groups to estimate TAIEX. The main input factors are previous Taiwan Stock, Dow Jones industrial average index, Nasdaq index, MIB, or their combination. First of all, the recommended approach is to transform the historical training data documents into the relationship between fuzzy groups and fuzzy logic. Secondly, it is formed according to the fuzzy variation groups of the minor factors to make fuzzy logical relationship groups(FLRGs). Thirdly, the reason why the major and minor factors are valued is to create the leverage between fuzzy variation and the fuzzy variation groups. Fourthly, it then appears in the data of each fuzzy variation of the fuzzy variation groups. Fifthly, the weighted calculation of the statistical data of fuzzy variations appears respectively in every fuzzy variation group [Chen 2011].

[Huang 2006] application SVN gray theory to predict the investment portfolio in the three years from 2003 to 2005 on Taiwan's stock market and Taiwan shares, via the SSVM investment in three years the average return rate of 306.386%, and the average return of 17.04% for the broader market display SSVM can effectively improve the rate of return on investment in the stock market. [Hsieh 2011] proposed the mixed computer wisdom mixed bee colony-recurrent NN method to predict the stock [Wu 2012] also proposed using SOPNN way combines hurt computer the wisdom speech compile method to predict the stock market and RMSE to calculate the results, good or bad, they are very good results. [Chen 2010] using Fuzzy-Trend with methods such as time series to forecast the Taiwan stock market closing price and test the data.

3. The Forecast Model

3.1 The E*xperimental Procedure*

In this study, historical data of main factors (i.e., weighted index of the Taiwan stock market) and historical data of secondary factors (i.e.,Dow Jones index, Nasdaq index, and M1B number in Taiwan) are gathered first. Then the predictive values of D-day are produced as:

* M1B number of D-day = M1B number of previous day,
* Taiwan stock index of D-day = Taiwan stock index of previous day * (1 + average fluctuation percentages of Dow Jones index and Nasdaq index in previous day),

Third, the historical data and predictive values are collected as input data of BPN algorithm. The forecast data are then produced by performing the BPN program. The assessment are done in the last step by using methods of average daily error, root mean square error (RMSE), trend ratio of fluctuation, and accuracy of fluctuation level.

3.2 Historical Data Gathering

Considering the annual line of average share index playing an important role in technical analysis of stock market, we take the last 12 months of historical data as the training data of the BPN algorithm. An another advantage is to eliminate the season's interference factors. In addition, considering the specific days (e.g., the Lunar New Year holidays, ex-dividend peak season in April and/or May, QFII's summer vacation in July and August) in the Taiwan stock market, we select, after repeated tries and tests, November and December to be the target time period of forecast. Its last 12 months of historical data are taken as the training data

To fit the BPN program, the following pre-process on historical data are necessary:
(a) Only the trading days of all of Taiwan, Dow Jones, and Nasdaq stock markets being open are selected. (b) The close time of Dow Jones and Nasdaq stock markets previous day is only five hours prior to the open time of Taiwan stock market current day. It results that Taiwan's investors often take the fluctuation of Dow Jones and NASDAQ markets previous day as the trend of Taiwan market current day. (D-1)-day index values of Dow Jones and Nasdaq markets are thus considered as D-day referent values of Taiwan market. That is, data of Dow Jones and Nasdaq markets on (D-1)-day are put on the field of that on D-day.

The sample data is shown in Table 1 after pre-processing, where SN means serial number, "T" in Marker field marks as training data, P as predict data. The TWSE indicates index values of Taiwan stock market. Data in TWSE column are taken as output of BPN program, while data in I-TWSE column as input.

3.3 Four Forecast Modes

There are four modes in predicting the Taiwan stock market in this study. (1) Use BPN approach referring to the prediction value calculated from fluctuation magnitudes of U.S. stocks market. (2) Use BPN approach directly. (3) Use BPN approach referring to the prediction value calculated from weighted moving average indexes of last ten days in Taiwan stocks market. (4) Use the prediction value directly calculated from fluctuation magnitude of U.S. stocks market.

3.3.1 Using BPN approach referring to U.S. market

Considering many Taiwan's investors often take the fluctuation magnitudes of Dow Jones and NASDAQ markets previous day as the trend of Taiwan market current day since the close time of Dow Jones and Nasdaq stock markets previous day is only five hours prior to the open time of Taiwan stock market current day. Based on this consideration, the prediction value of Taiwan market is then calculated from fluctuation magnitude of Dow Jones and Nasdaq stocks markets previous day. That is, the prediction value of Taiwan stock market is calculated as:
predication index of D-day = index value of previous day * (1 + average fluctuation percentages of Dow Jones index and Nasdaq index in previous day), [eq.1]

For example, the predication index value of I-TWSE column on 2009/11/01 in Table 1 is calculated as follows:

6957.40 * (1+ (((8975.46/8915.47)-1)+ ((1856.56/1837.10)-1)))/2) = 7017.65

The data in Table 1 is then fed into BPN program to calculate the forecast value of TWSE column on 2009/11/01 in Table 1.

3.3.2 Using BPN approach directly

Choose data from Marker column to M1B column in Table 1 and delete the I-TWSE column to build a new table. Take this new table as input of BPN program. It is called traditional BPN approach.

3.3.3 Using BPN approach referring to weighted indexes of Taiwan market.

We consider the index value of Taiwan stock market is major affected by its indexes in previous ten days in this mode. Weighted moving average method is used to derive an index value based on the variously affecting power of index of each day. Different weight is appended to index of various days. In common sense, the nearer day is more affecting. We calculate the prediction value by using linear function on data of previous ten days.

For example, assume that the index values of Taiwan stock market are 6790.15, 6833.26, 6822.13, 6879.73, 6852.06, 6920.44, 6852.39, 6903.84, 6889.65, and 6957.40 respectively in sequence at the previous ten trading days before 2009/11/01. The predication index value of I-TWSE column on 2009/11/01 in Table 1, in this mode, is calculated as follows:

(10/55*6957.40)+(9/55*6889.65)+(8/55*6903.84)+(7/55*6852.39)+(6/55*6920.44)+(5/55*6852.06)+(4/55*6879.73)+(3/55*6822.13)+(2/55*6833.26)+(1/55*6790.15) = 6890.97.

3.3.4 Using the prediction directly from U.S. stocks market.

In this mode, we take the calculated value from fluctuation magnitude of Dow Jones and Nasdaq markets previous day, just like that shown in eq.1, as the forecast index value of Taiwan stock market current day. This approach is common used by most stock market guru named TV channels mouth.

3.4 Feeding input data into BPN program

Please note that, except the approach of directly calculation from U.S. stocks market, which is shown in section 3.3.4, the input data should be fed into the BPN program to train and to get the forecast value for the other three modes.

3.5 Performance Indicators

Daily Average Error (DAE) and Root-Mean-Square-Error (RMSE) are shown in eq.2 and eq.3 respectively.

Accuracy Judgement for Forecast Values

* Change the actual value = actual closing price of the actual closing price of the D-day - (D-1)-day.
* Estimated Change value = D-day Estimate - (D-1)-day closing price.
* Forecast stock price movements accurately value = 1 if (Estimate value / Actual value)> = 0, 0 if (Estimate value / Actual value) <0.
* Estimated Change rate accurately value = (Estimated values - Actual value) / Actual value.
* Forecast accuracy (%) = Estimated rate x accurately forecast stock price movements accurately value.

4. Experimental Results and Evaluation

4.1 Forecast Accuracy Related to the Amount of Training Data

Our study shows that, the forecast accuracy using BPN algorithm is independent from the amount of training data. Take an example, to predict the stock index values of coming five days in one program run, selecting test data of two years randomly (They are 2004 and 2006 here), different amount of training data, i.e., data of two months, four months, six months, eight months, ten months, and twelve months respectively are provided. The experimental results are shown in Table 2.

Another example to take training data of 10 months and 12 months respectively to predict the stock index values of coming two months in one program run, 12 years (1999 to 2010) test data are calculated. The experimental results are shown in Table 3.

The above experimental results show that, the forecast accuracy using BPN algorithm is independent from the amount of training data.

4.2 Accuracy Comparison for Short-term and Long-term Forecasts

We tried to compare the forecast accuracy between short-term forecast and long-term forecast. We used training data of last 10 months to predict the stock index values of coming five days, ten days, twenty days, thirty days, and sixty days after Nov. 1 respectively in one program run, 12 years (1999 to 2010) test data are calculated. The experimental results are shown in Table 4.

The above experimental data show that, short-term forecast is more accuracy than long-term forecast in the condition of using BPN algorithm with the same amount of training data. And the forecast is longer term, the accuracy is worse.

4.3 Accuracy Comparison for Forecast One-day and Forecast Five-days Once

We then compare the forecast accuracy between forecasting one-day once and forecasting five-days once. Forecasting five-days once forecasts the index values of coming five days in one program run, while forecasting one-day once forecasts the index values of coming day in one program run and perform the program five times for sequential five days. The same training data of last 12 months are used. 12 years (1999 to 2010) test data are calculated. The experimental results are shown in Table 5.

The above experimental data show that, forecasting one-day once is more accuracy than forecasting five-days once in the same condition.

5. Concluding Remarks

In this study, the combination of wisdom and humanity expectations applied to the prediction of the stock market, is different from the existing literature on the use of a single way to predict the stock market. This study proposes a the four prediction method, the six-year real comparison, and made a rigorous estimate the accuracy of authentication for price estimate another direction of thinking, this is a contribution of this study to provide investors.

Acknowledgment

The authors thanks National Science Council, Taiwan R.O.C. and Tatung University, Taiwan R.O.C. for their kindly financial supports under the grand contracts NSC101-2221-E-036-019 and B101-I04-036 respectively.

REFERENCES

[Wu 2012] Jui-Yu Wu and Chi-Jie Lu "Computational Intelligence Approaches for Stock Price Forecasting" International Symposium on Computer, Consumer and Control, Taichung, Taiwan, Jun 4 2012, pp. 52 - 55.

[Lin 2012] Cheng-Ting Lin and Li-Der Chou ,"Incorporating Stock Index in a Support Vector Regression Model to Improve Short Term Load Forecasting Accuracy", International Symposium on Computer, Consumer and Control, pp.634-637.

[Hsieh 2011] T. J. Hsieh, H. F. Hsiao and W. C. Yeh, "Forecasting Stock Markets Using Wavelet Transforms and Recurrent Neural Networks: An Integrated System Based on Artificial Bee Colony Algorithm," *AppliedSoft Computing*, vol. 11, no. 2, 2011, pp. 2510-2525.

[Chen 2011] Shyi-Ming Chen and Chao-Dian Chen, "TAIEX Forecasting Based on Fuzzy Time Series and Fuzzy Variation Groups"， IEEE Transactions on Fuzzy Systems, Vol. 19, No. 1, February 2011.

[Chen 2010] Shyi-Ming Chen and Nai-Yi Wang, "Fuzzy Forecasting Based on Fuzzy-Trend Logical Relationship Groups", IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, Vol. 40, No. 5, October 2010.

[Wong 2010] Wai-Keung Wong, Enjian Bai, and Alice Wai-Ching Chu, "Adaptive Time-Variant Models for Fuzzy-Time-Series Forecasting", IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics, Vol. 40, No. 6, December 2010.

[Kennedy 1995] Kennedy J. and Eberhart R. C., "Particle swarm optimization,", Proc. IEEE Int'l. Conf. on Neural Networks, Vol.4, Piscataway, NJ: IEEE Service Center, 1995, pp. 1942-1948.

[Chen 2009] C. D. Chen and S. M. Chen, "A new method to forecast the TAIEX based on fuzzy time series," in Proc. 2009 IEEE Int. Conf. Syst., Man, Cybern., San Antonio, TX, 2009, pp. 3550–3555.

[Chen 1996] S. M. Chen, "Forecasting enrollments based on fuzzy time series," Fuzzy Sets Syst., Vol. 81, No. 3, 1996, pp. 311–319.

[Chen 2002] S. M. Chen, "Forecasting enrollments based on high-order fuzzy time series," Cybern. Syst., Vol. 33, No. 1, pp. 1–16, 2002.

[Chen 2006] S. M. Chen and N. Y. Chung, "Forecasting enrollments using high-order fuzzy time series and genetic algorithm," Int. J. Intell. Syst., Vol. 21, No. 5, 2006, pp. 485–501.

[TAIEX Web Site] TAIEX Web Site, http://www.twse.com.tw/en/products/indices/tsec/taiex.php, visited on Jun. 21, 2012.

I. [Zurada 1994] J.M. Zurada, "Sensitivity analysis for minimization of input data dimension for feedforward neural network", 1994 IEEE International Symposium on Circuits and Systems, May 30 1994, London, U.K., pp.447-450 of vol. 6.

II. [Hashem 1992] S. Hashem, "Sensitivity analysis for feed-forward artificial neural networks with differentiable activation functions", International Joint Conference on Neural Networks, Jun 7 1992, Baltimore, MD, USA, pp.419 - 424.

[Ghosh 2008] S. Ghosh-Dastidar, "Principal Component Analysis-Enhanced Cosine Radial Basis Function Neural Network for Robust Epilepsy and Seizure Detection", IEEE Transactions on Biomedical Engineering, Feb. 2008, Vol. 55(2), pp.512 - 518.

[Gevrey 2003] Muriel Gevrey, Ioannis Dimopoulos, and Sovan Lek, "Review and comparison of methods to study the contribution of variables in artificial neural network models", Ecological Modelling, Elsevier B.V., U.K., Vol. 160(3), Feb. 15 2003, pp. 249–264.

[Gevrey 2006] Muriel Gevrey, Ioannis Dimopoulos, and Sovan Lek, " Two-way interaction of input variables in the sensitivity analysis of neural network models", Ecological Modelling, Elsevier B.V., U.K., Vol. 195(1), May 15 2006, pp. 43–50.

[Armano 2005] G. Armano, M. Marchesi, and A. Murru, "A hybrid genetic-neural architecture for stock indexes forecasting", Information Sciences, Elsevier B.V., U.K., Vol. 170(1), Feb. 18 2005, pp. 3–33.

III. [Atsalakis 2009] George S. Atsalakis and Kimon P. Valavanis, "Surveying stock market forecasting techniques – Part II: Soft computing methods", Expert Systems with Applications, Elsevier B.V., U.K., Vol. 36, 2009, pp.5932–5941.

IV. [Hadavandi 2010] Esmaeil Hadavandi, Hassan Shavandi, Arash Ghanbari, "Integration of genetic fuzzy systems and artificial neural networks for stock price forecasting", Knowledge-Based Systems, Elsevier B.V., U.K., 2010, doi:10.1016/j.knosys.2010.05.004.

[Shao 2012] Xigao Shao, Kun Wu, and Bifeng Liao, "l_p-Norm Multikernel Learning Approach for Stock Market Price Forecasting", Computational Intelligence and Neuroscience, Hindawi Publishing Corp., Egypt, Vol. 2012, Article ID 601296, 10 pages, doi:10.1155/2012/601296

[Tsang 2009] Edward Tsang, "Forecasting — where computational intelligence meets the stock market", Frontiers of Computer Science in China, Springer Publishing, USA, Volume 3(1), Mar. 2009, pp. 53-63.

[Kim 2006] Kyoung-jae Kim, "Artificial neural networks with evolutionary instance selection for financial forecasting", Expert Systems with Applications, Elsevier B.V., U.K., Vol. 30(3), April 2006, pp. 519–526.

About Authors:

Jin-Cherng Lin received the Ph.D. degree in Information Engineering and Computer Science from National Chiao-Tung University, Taiwan, in 1989. He is currently an associate professor in the Department of Computer Science and Engineering at Tatung University, Taiwan. His research interests include software testing and validation, software quality assurance, computer network management, and computer network security. His email is jclin@ttu.edu.tw.

Fu-Jyh Ko and **Jian-Hon Liu,** received their M.S. degrees in Information Engineering and Computer Science from Tatung University, Taiwan, in 2012. They are currently R&D engineers. Their research interests include computing intelligence techniques, software engineering, and system design.

Yueh-Ting Lin received his M.S. degree in Mechanical and Electro-Mechanical Engineering from National Sun Yat-Sen University, Taiwan, in 2011. He is currently an R&D engineer. His research interests include computing intelligence techniques, intelligent transformation system, and digital signal process.

$$\text{DAE} = \sum_{i=1}^{n} \text{ABS}\left(\text{forecasted value}_i - \text{actual value}_i\right)/\mathbf{n} \qquad \text{(eq.2)}$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}\left(\text{forecasted value}_i - \text{actual value}_i\right)^2}{n}} \qquad \text{(eq.3)}$$

Table 1: Sample Data after Pre-processing

date	SN	Marker	TWSE	I- TWSE	Dow Jones	Nasdaq	M1B
2008/11/01	1	T	7218.09	7218.09	8592.10	1771.39	3644548
2008/11/02	2	T	7071.44	7071.44	8706.15	1800.91	3654181
2008/11/03	3	T	6905.32	6905.32	8706.15	1788.43	3649241
…………	…	T	………	………	………..	………	……….
2009/10/30	240	T	6889.65	6889.65	8922.85	1836.49	3666535
2009/10/31	241	T	6957.40	6957.40	8915.47	1837.10	3664467
2009/11/01		P		7017.65	8975.46	1856.56	3664467

Table 2. Forecast Accuracy Related to the various amount of training data.

Year	Error Index	Two Months	Four Months	Six Months	Eight Months	Ten Months	Twelve Month	Aver-age	Standard Deviation
2004	DAE	73.28	73.50	73.40	73.06	73.22	73.21	73.28	0.14
2004	RMSE	102.75	100.96	101.50	102.10	101.59	101.78	101.78	0.55
2006	DAE	49.13	47.68	48.74	47.89	48.20	48.04	48.28	0.50
2006	RMSE	58.24	55.68	60.19	57.29	58.05	57.82	57.88	1.34

Table 3(a). Experimental Forecast Accuracy for the 10 months of training data.

Error Index	1999	2000	2001	2002	2003	2004	
DAE	136.29	192.56	306.72	109.95	125.38	55.40	
RMSE	184.18	249.78	349.87	142.92	164.44	73.72	
Error Index	2005	2006	2007	2008	2009	2010	Average
DAE	68.03	110.38	286.28	225.78	105.41	241.33	163.63
RMSE	83.02	140.64	346.14	265.86	145.78	311.31	204.81

Table 3(b). Experimental Forecast Accuracy for the 12 months of training data.

Error Index	1999	2000	2001	2002	2003	2004	
DAE	146.91	190.88	216.72	178.97	133.46	54.70	
RMSE	186.03	244.37	275.03	218.60	162.23	72.50	
Error Index	2005	2006	2007	2008	2009	2010	Average
DAE	83.26	135.71	316.31	190.80	108.85	239.83	**166.37**
RMSE	99.52	155.59	389.70	228.55	152.34	311.07	**207.96**

Table 4. Comparison of Forecast Errors for Short-term and Long-term Forecasts.

Forecast Period	5 Days	10 Days	20 Days	30 Days	60 Days	Comparison
DAE	140.87	183.14	257.05	327.39	391.33	Short-term is more accuracy than long-term.
RMSE	158.20	207.27	293.76	374.10	449.62	

Table 5. Accuracy Comparison for Forecast One-day and Forecast Five-days Once.

Forecast type	Forecast Five-days Once	Forecast One-day Once
DAE	139.62	132.67
RMSE	157.62	153.53