International Journal of Engineering Science Invention
ISSN (Online): 2319 – 6734, ISSN (Print): 2319 – 6726
www.ijesi.org Volume 3 Issue 5‖ May 2014 ‖ PP.36-39

Stokes Trading in Market using Neural Network

Kavita Saini[1], Abhishek Upadhyay[2]
M.Tech.Scholar,[1] Asst. Professor[2]
[1,2,]Electronics and Communication Department CBS Group of Institution, Jhajjar Haryana

ABSTRACT :*The main focus of this study is to compare different performances of soft computing paradigms for predicting the direction of individuals stocks. Three different artificial intelligence techniques were used to predict the direction of both Microsoft and Intel stock prices over a period of thirteen years. We explored the performance of artificial neural networks trained using back propagation and conjugate gradient algorithm and a Mamdani and Takagi Sugeno Fuzzy inference system learned using neural learning and genetic algorithm. Once all the different models were built the last part of the experiment was to determine how much profit can be made using these methods versus a simple buy and hold technique.*

KEYWORDS: *Neural Network, linguistic variables, stock prices.*

I. INTRODUCTION

The ability to predict the direction of the stock prices is the most important factor to making money using financial prediction. All the investor really needs to know is to buy if the stock is going up in value and to sell if it is decreasing in value. This paper will delve into some of the most popular soft computing techniques for stock market modeling [1]. These methods are: neural networks, fuzzy inference system, genetic algorithm, and some heuristic techniques. The most recent studies compare indexes such as the S&P 500, NASDAQ, and the Dow Jones[2][4][5][8][10]. The experiments done in this project examine the chaotic behavior of actual company stocks that tend to be less stable and thus harder to predict. Studies have also shown that using direction as compared to prediction can generate higher profits [4], and this study will try and capitalize on that idea. Also the prediction will examine a more realistic situation where an investor has the choice between multiple stocks, in this case 2, and chooses the stock that is mostly likely to increase in value.

The experiments also compare many hybrid AI techniques and their abilities to predict acategorical output. The data for this research comprised of prices for Microsoft and Intel Corporations from January 2nd, 1990 until August 5th, 2003. Information contained for each daily report is the opening price, closing price, low price, high prices, and the volume of shares traded. Economic indicators such as the current Prime rate, Michigan's Consumer Sentiment Index, and the United States Consumer Confidences were also used to aid the models during the training phase. We explored the performance of artificial neural networks trained using back propagation and conjugate gradient algorithm and Mamdani and Takagi Sugeno Fuzzy inference system learned using neural learning and genetic algorithm for directional prediction.

II. RELATED RESEARCH

Many papers have dealt with input selection when it comes to mapping financial indexes and stocks [2][5][16][18][19]. Inputs have been broken into two different types of inputs, financial and political (which tend to be qualitative). Kuo et al. [7] uses a genetic algorithm base fuzzy neural network to measure the qualitative effects on the stock price. Variable selection is critical to the success of any network for the financial viability of a company. Quah et al. [9] identified 5 key parts namely yield, liquidity, risk, growth, and momentum factors. Macroeconomic factors such as inflation and short-term interest rate [5] have to shown to have direct impacts on the stock returns. A better measure of fitness, which considers profit [10], has been suggested to replace a root means squared error. Yao and Poh [12] showed an example where a model with a low Normalized Mean Square Error (NMSE) had a lower return than a model with a higher NMSE. Brownstone [3] recommends using percentages to measure performance so that the result can be better understood by traders and other people that might need their research and are not experts in the field. Chen et al. [4] used a sliding window to predict the next day's price of the index. Every day the network was retrained with the most recent 68 days of input with the attempt to predict the coming day. Commission (remuneration for services rendered) is commonly overlooked when doing research relating to stock market prediction; however, if any model is actually implemented it is going to incur fees which could greatly affect the profit predicted by the model.

Chen et al. [4] considers 3 different levels of commissions and how it would affect the best buying strategy used by investors.

III. HURST EXPONENT

Some papers have used the Hurst Exponent [10] to prove that the data is not completely random but in fact has the correspondence between the input and the output data. The Hurst Exponent can show the degree of correlation. If the exponent is 0.5 the data is completely random and no thus no network will be able to predict the output and thus it is a waste of time to attempt to learn any pattern in the data. The closer the Hurst Exponent is to one, the greater the correlation between the input and output, and a Hurst Exponent of less than 0.5 means that the input and output are indirectly proportional. It is important to note the Hurst Exponent is confined to the range of 0 to 1.

$$Hurst\ Exponent = \frac{\log\left(\frac{R}{S}\right)}{\log\left(S\right)} \qquad (3.1)$$

Fuzzy logic gives a set of natural language rules that are easily understood by humans. The primary advantage of fuzzy logic is its readability. For a first order Takagi-Sugeno model, a common rule is represented as [14]:

If x is A_1 and y is B_1, **then** $f_1 = p_1x + q_1y + r_1$ (3.2)

where x and y are linguistic variables and A_1 and B_1 are corresponding fuzzy sets and p_1, q_1, r_1 are linear parameters. Usually the least squares algorithm is used to determine the linear parameters and the membership function parameters are fine tuned using a neural network learning method. Initial rules are generated using the grid-partitioning method [6]. In the inference method proposed by Mamdani the rule consequence is defined by fuzzy sets and has the following structure [13]:

if x is A_1 and y is B_1 then $z_1=c_1$ (3.3)

where x and y are linguistic variables, A1 and B1 are input fuzzy sets and C1 the corresponding output fuzzy set. Usually a mixture of neural learning and global optimization method are used to fine tune the various rule parameters. Genetic algorithms loosely mimic the concept of natural selection. Each member is made up of a chromosome, which is normally a binary string. This chromosome defines the characteristic of the member of the population and that allows the algorithm to determine its fitness. A population is a group of members and changes from generation to generation through methods S is the standard deviation of the time series before normalization and R is the maximum and minimum cumulative deviations of the observation has compared with the mean of the series. N is the number of observations

$$R_N = \max_{1\leq t\leq N}[x_{t,N}] - \min_{1\leq t\leq N}[x_{t,N}] \qquad (3.4)$$

xt ,n , the cumulative deviation, is describe by

$$x_{t,N} = \int_{1}^{t} xut\ dt\ \ for\ all\ N \qquad (3.5)$$

Where x is the mean of xu for all N elements. The Hurst exponent can be very useful in any set and allows a method of comparing sets of data.

IV. SOFT COMPUTING

Soft computing comprises of new generation computationally intelligent hybrid systems consisting of neural networks, fuzzy inference system, approximate reasoning and derivative free optimization techniques. Neural Networks is an attempt at creating a computer that could learn in a manner similar to humans. Neural Networks can determine complex function approximations, classifications, auto associations etc. such as mutation and crossover. The fitness function is used at every generation to see which members are fit and most likely to survive to the next generation through a reproduction process.

V. RESULTS DISCUSSION

The data sets used in this paper originally had data from January 1990 to August 2003. Since the stock prices had a big variation during the entire period, all the data prior to 1997 was removed from the data set. This provided 2 services: first it reduced the size of the data set and second it gave a better split of increasing days to decreasing days. Studies have shown in classification problems it is important to have equal representations of both cases in order to prevent the network from becoming biased towards the more common value, in this case increasing days. Intel data set contained less than 1% difference in the representation of increasing days as compared with decreasing days, thus further manipulation of the data set was not required. However Microsoft data during this period had an increase in 59.6% of the days in the sample set. In order to prevent the network from heavily favoring the increased prediction, the data set had increasing predictions randomly removed until a 55/45 split was achieved. This allowed experiments to be tested to show the difference with strong bias and

without a bias. Thus the last year of data from August 1st, 2002 until July 31st, 2003 was held out of set to be used as an unseen testing set in the simulation. All days in which the price increased or decreased by more than 10% in a single day were removed from the database as these outliers were most likely caused by external forces. Also days in which less than 0.1%, or no change occurred were removed because an action performed by the investor would not affect the bottom line. The outliers for these models were determined to be days when trade volume is 4 times the average trade volume or more. For Microsoft, this value came out to any day that more than 105 million shares were traded. The prime rate [11] was also used and any day that rate was changed was removed from the data set. Transforming the data into a network usable form is essential for the success of the network. The data in this paper was normalized using min-max normalization. There are 16 inputs to begin with for the two data sets. The original 16 inputs that were considered are: consumer confidence index, the prime rate, Michigan consumer sentiment Index, price -1 (yesterday), price -2 (day before yesterday), price -3, price -4, price -5, volume -1 (volume yesterday), low -1, high -1, open -1, low -2, high -2, open -1, and volume -2. Several models were used to determine the least important inputs and then these inputs were removed before a more accurate and iterative approach could be used to determine the final inputs. A genetic algorithm with a population size of 50 was trained for 100 generations and the sensitivity about the mean test one input randomly. The network which did the best on test was kept, thus after the first iteration there were 8 inputs left. All networks were trained 3 times with randomly initialized weights and the best network (selection of input variables) was chosen. The neural network used 20 and 7 neurons in the hidden layers. The conjugate gradient learning algorithm was used for 10,000 epochs for all the experiments. Table 1 shows the results that determined CCI (Consumer Confidence Index) should be removed from Microsoft data. The network with CCI input had a NMSE of 0.007506, which was worse than the network that had the input removed (0.007119). elects volume to be removed with a NMSE of 0.006871.

Input variable	MSE cross validation	NMSE on testing
CCI	0.000520791	**0.007118908**
close-3	0.000535407	0.007293936
volume	0.000527196	0.007308423
high-2	0.000514665	0.007397344
close-2	0.000529705	0.007571371
open-1	0.000515599	0.007645692
high-1	**0.000512655**	0.007789516
close-1	0.000600180	0.008804343

Table 1. Snapshot of input reduction with 8 variables

Input variable	MSE cross validation	NMSE on testing
volume-1	0.000514894	**0.006871094**
close-2	0.000517851	0.007191489
high-2	**0.000509082**	0.007259947
close-3	0.000519643	0.007460574
high-1	0.000532197	0.007658685
open-	0.000524406	0.007685916
close-1	0.000583583	0.008093149

Table 2. Snapshot of input reduction with 7 variables

This process was iterated until the dataset contained the 5 most significant inputs. The final selection of inputs is: close-1, open-1, high-1, high-2, and close-3. A soft computing paradigm could be used to model a regression or classification task. When performing regression, all the 5 inputs are given to the network, and the output is the stock price for the next day. Once a test is completed the predicted price is compared with yesterday's price to see if the price increased or decreased. Classification is much more straightforward. Before presenting the data to the network the output is translated to 1, for increase, or 0 for decrease. And the objective of the network is to correctly predicted 0 or 1.

VI. CONCLUSIONS

The ability to predict stocks on a daily basis is a very difficult problem even for the most advance networks. The Hurst component confirmed the hypothesis, which is prediction is possible but especially difficult. Surprisingly, the Consumer Confidence Index and Prime Rate were not able to improve the predictability of these networks. It is clear from all the tests that the networks were able to learn the pattern in Microsoft's data much more easily than Intel's data, thus it might be possible that another stock is more learnable than Microsoft. Through the uses of many techniques it is possible to correctly predict the direction of the stock 63% of the time for a large company like Microsoft. Thus, this research provides the groundwork for financial trading using some of the well known soft computing paradigms. Even the worst model used for Microsoft produced a return on investment of 66%, and the best network scored an astounding 103% return. This study also demonstrated that picking the correct stock is as important as building the best network; as the best network for Intel was outperformed by the worst network on Microsoft data. The biggest downfall of these networks is that the transaction cost of buying and selling stocks would be very costly. However, it would be feasible to fine tune the buy and sell strategy to lower this cost.

REFERENCES

[1] Abraham A., Intelligent Systems: Architectures and Perspectives, Recent Advances in Intelligent Paradigms and Applications, Abraham A., Jain L. and Kacprzyk J. (Eds.), Studies in Fuzziness and Soft Computing, Springer Verlag Germany, Chapter 1, pp. 1-35, 2002.

[2] Abraham, A., Philip, N. S., and Saratchandran, P. "Modeling Chaotic Behavior of Stock Indices Using Intelligent Paradigms.Neural, Parallel and Scientific Computations, 11 (2003): 143-160.

[3] Brownstone, D. Using Percentage Accuracy to Measure Neural Network Predictions in Stock Market Movements. Neurocomputing 10(1996): 237-250.

[4] Chen, A.S., Leung, M.T., and Daouk, H. Application of Neural Networks to an Emerging Financial Market: Forecasting and Trading the Taiwan Stock Index. Computers and Operations Research 30 (2003): 901-923.

[5] Izumi, K. and Ueda, K. "Analysis of Exchange Rate Scenarios Using an Artificial Market Approach." Proceeding of the International Conference on Artificial Intelligence 2 (1999): 360-366.

[6] Jang, J.S.R. "ANFIS: Adaptive-Network-Based Fuzzy Inference System." IEEE Transactions on Systems, Man, and Cybernetics 23 (1993): 665-684.

[7] Kuo, R.J., Chen, C.H., and Hwang, Y.C. "An Intelligent Stock Trading Decision Support System through Integration of Genetic Algorithm Based Fuzzy Neural Network and Artificial Neural Network." Fuzzy Sets and Systems, 118 (2001): 21- 24.

[8] O'Brian, T. V. "Neural Nets for Direct Marketers" Marketing Research, Volume 6, Issue 1

[9] Quah, T.S. and Srinivasan, B. "Improving Returns on Stock Investment through Neural Network Selection." Expert Systems with Applications 17 (1999): 295-301.

[10] Yao, J.T. and Tan, C.L. "A Study on Training Criteria for Financial Time Series Forecasting." Proceedings of International Conference on Neural Information Processing. Nov. 2001: 772-777.

[11] Prime rate: http://research.stlouisfed.org/fred2/data/PRIME.txt

[12] Yao, J. and Poh, H.L. "Forecasting the KLSE Index Using Neural Networks." IEEE International Conference on Neural Networks 2 (1995) 1012- 1017.

[13] Mamdani E H and Assilian S, An experiment in Linguistic Synthesis with a Fuzzy Logic Controller, International Journal of Man-Machine Studies, Vol. 7, No.1, pp. 1-13, 1975.

[14] Sugeno M, Industrial Applications of Fuzzy Control, Elsevier Science Pub Co., 1985.