RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS No. 104
Business Informatics 16 2010

Maciej Bac
Wrocław University of Economics, Wrocław, Poland
maciej.bac@gmail.com

SELF ORGANIZING MAP (SOM) NETWORK APPLICATION SUPPORT FOR SHORT-TERM INVESTMENT DECISIONS

Abstract: Popularity of global futures markets grows every year. Attracted by the possibility of quick and easy profits, investors increase liquidity of the base instruments and derivatives. Liquidity growth brings a feeling of security and attraction to markets. Professional investors introduce more and more advanced techniques and tools to support analysis of financial instruments. Computers are used not only to perform standard analysis but also to utilize artificial intelligence as a tool for prediction. This is an area where much more remains to be discovered than has already been defined, and hence the author's interest in this topic. Referring to those observations, the objective of this article is to present (and review) the concept of SOM neural network application for investment decisions support. The author has presented results of his own research based on the developed system, which is a practical attempt to verify the presented idea.

Keywords: neural networks, stock exchange, SOM.

1. Introduction

Popularity of global futures markets grows every year. Attracted by the possibility of quick and easy profits, investors increase liquidity of the base instruments and derivatives. Liquidity growth brings a feeling of security and attraction to markets. Professional investors introduce more and more advanced techniques and tools to support analysis of financial instruments. Investments funds recruit to their group of brokers and financial analysts also computing scientists. They are supposed to help increase investment returns by predicting trends, economic cycles and changes in the markets. Computers are used not only to perform standard analysis but also to utilize artificial intelligence as a tool for prediction. This is an area where much more remains to be discovered than has already been defined, and hence the author's interest in this topic.

Traditional methods of analyzing and forecasting do not produce satisfactory results. Appropriate use of genetic algorithms, which usually gives good results for searching through vast space of solutions in a short time, could be a breakthrough in the analysis of stock markets [Bac, Kwaśnicka 2009]. Neural networks, with their excellent ability to learn and generalize, could be able to solve the problem of predicting stock market behavior. The combination of investments signals taken from methods mentioned above combined into multi-agent system [Korczak, Lipiński 2008, 2006] allows us to do more secure placement of money in financial markets.

Referring to those observations, the objective of this article is to present (and review) the concept of self organizing map (SOM) neural network application for investment decisions support. First of all the author has explained the base assumptions of his idea, then presented a model of its implementation in a computer system. At the end, the author has presented the results of the research, conducted by the author and based on the developed system, which is a practical attempt to verify the presented idea.

2. Base assumptions of the idea

The idea described in this article is based on three fundamental concepts which have been explained and described in this part. These characteristics include technical analysis, Japanese candles and SOM neural networks.

Technical analysis and fundamental analysis play an important role in the analysis of the stock market. The purpose of fundamental analysis is to monitor and classify financial assets in terms of their investment grade (as an estimate of the risk assessment) and the expected return of investment [Ritchie 1997]. Fundamental analysts estimate the value of company, if they consider that the company is undervalued according to the market, it is recommend to buy, otherwise – for sale. Analysts consider such factors as: current and projected macroeconomic situation of the region and country, interest rates, industry growth opportunities, etc. Necessary information to generate the report and describe the status of the company and its prospects for long and medium term is not easily accessible for typical shareholders. Therefore, the subject of interests of the author is technical analysis. Technical analysis is based on a study of the effects of market behavior (price, volume) and finding on the price charts trends and formations (patterns) to help predict market behavior [Ikeda, Tokinaga 2007; Li, Tsang 1999]. Stock market analysts often recognize different formations in the same period of time. The formation may extend over a long or short term, any of the phases of the development of the formation may not occur, so it is difficult to describe them algorithmically. In contrast, Japanese candles are precise, and therefore they were selected for analysis [Bac, Kwaśnicka 2009].

Candle charts are based on Japanese candles. They are one of the oldest methods of charts analysis, developed in Japan over 300 years ago. To Europe and the United States arrived in the 1990s, quite well adopted also in Poland. To construct Japanese candle required values are opening, closing, maximum and minimum price of the period. Daily candle provides information on the opening, closing, maximum and minimum price for that day. The length of the candle symbolizes the difference between the maximum and minimum price of the period. By contrast, candles body length symbolizes the difference between the opening and closing price. Candle color says, if the price has increased since the opening of the time period or dropped. White candle symbolizes the period of growth, while the black candle represents a fall. Depending on the period that candle symbolizes it represents long, medium or short term.

SOM neural networks have been frequently used to support investment decisions. They are perfectly suitable for detecting similarity between technical or fundamental situation of the individual stock in the financial markets. Kohonen's network based on fundamental analysis indicators and grouped values of stock exchanges in Mumbai [Khan, Bandopadhyaya, Sharm 2009] and U.S. [Stankevicius 2001] helps to optimize investment portfolio. SOM network supported matching stocks with similar charts in the past 30 days [Simunic 2003] what made easier to find the company with a similar technical situation. Neural network [Afolabi, Olude 2007], which analyzes 57 last quotations, the maximum and minimum price and the difference between the previous and the current closing price, is used to support the decision to purchase or sale stock. The possibility of using self organizing networks to find similarities between the current situation and historical situations inspired the author to propose the following solution.

3. Idea implementation model – system architecture

A general model of the system, in which the concept of SOM neural network application to support investment decisions has been implemented, is presented in Figure 1. Three of its base modules are characterized below and are the following:
- module clustering/pattern recognition,
- the statistical module,
- the decision-making module.

The principle of the system operation is as follows (Figure 1): information of changes in quotation, provided by the data transfer module, is clustered by the Kohonen's network and the winning neuron is returned. Decision-making module collects statistical data about the winning neuron from the analysis module. Finally, the system decides, whether open or close position, on the basis of statistics and the given parameters.

The clustering/pattern recognition module is based on SOM neural network and is implemented as a part of an application in *C#*. Neuron represents the *N* subsequent candles. We can decide whether to take into consideration data from open, close, high, low or average of open and closing. Figure 2 presents visualization of a neuron representing 30 subsequent candles – in the opening and closing values.

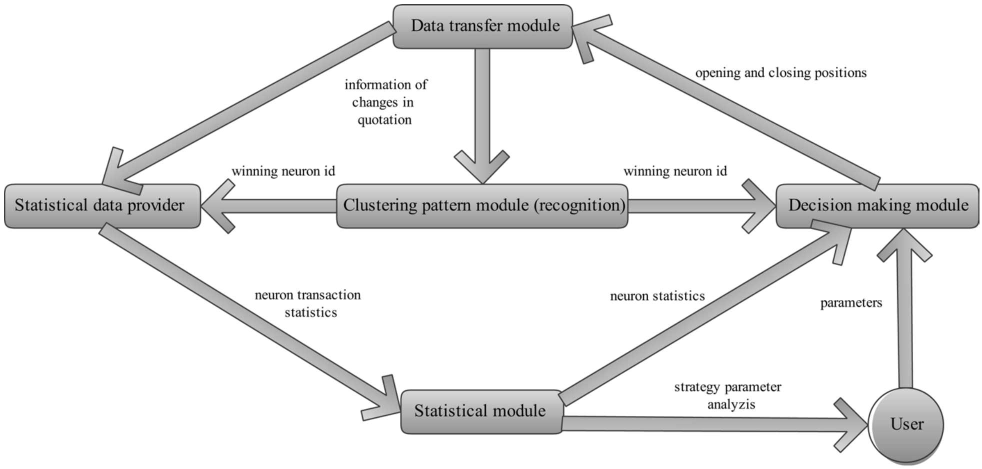

Figure 1. A general model of the system

Figure 2. Visualization of a neuron representing 30 subsequent candles – the opening and closing values

The input signal for such a network is the representation of current technical situation for selected stock. Representation must be the same as for the neuron from specific network, so it must take into consideration the recent *N* candles and candle parameters (opening, closing, minimum price, maximum price, etc.). Values inside the neuron and the input signal are normalized to 1. For each input the signal network returns neuron, which is most similar to the input signal. Depending on the quality of the network learning selected pattern is more or less similar to the input signal.

The statistical module, implemented in MS SQL database, stores data which describe the behavior of the quotation after the occurrence of a specified pattern/neuron on the chart. The range of time which is observed is a parameter. Other observed values are: the **maximum price** after the pattern occurrence, the **minimum price** after the pattern occurrence and **opening** and **closing price** after the time which is a parameter. Based on the data collected in this module, we can determine the parameters describing the neuron. These parameters include:

– the average maximum deviation up quotation AMDU (the average of maximum deviations of the quotation to the top of all periods opened by the neuron);
– the average maximum deviation down quotation AMDD (the average of maximum deviations of the quotation to the bottom of all periods opened by the neuron);
– the factor of opening position direction FDOP (the difference between the average of maximum deviation up quotation and the average maximum deviation down quotation). This ratio shows whether it is better to open a long or short position after activating the neuron (greater than zero – long, less than zero – short). The higher absolute value of ratio is, the stronger the signal to open position;
– deviation of the value of opening direction factor between the periods of research VFDOP;
– the number of positions which have been opened by the neuron;
– deviation of quantity of the opened positions by the neuron.

The above parameters help decide what to do, when the pattern recognition module will select the neuron which best suits the current situation of observed stock. Decision-making module can make decision when to invest based on the information which neuron is most suited to the current situation on the chart and information on the preferred direction of the quotation after that neuron activation in the historical periods. In this module three investment strategies have been implemented.

1. Basic strategy (Strategy 1) – a strategy that opens position in the direction determined by FDOP. The position is closed at the time which is a parameter constrained in advance. External parameters: the minimum value of the FDOP to open position, the maximum value of the coefficient of VFDOP to open position.

2. Strategy FDOP (Strategy 2) – a strategy that opens position in the direction determined by FDOP. Position is closed in two situations. First, if it reaches the expected value of the quotation, which is calculated using the formula:

$$EV = O + (O * \mathrm{FDOP} * M),$$

where: EV – expected value,
O – opening price,
FDOP – factor direction of opening position,
M – multiplier, external parameter.

Second, if the quotation does not achieve the expected value, the position is closed at the time which is a parameter constrained in advance. External parameters: the minimum value of the FDOP to open position, the maximum value of the coefficient of VFDOP to open position, the multiplier.

3. Strategy of the average deviation up/down (Strategy 3) – a strategy that opens position in the direction determined by FDOP. Position is closed in two situations. First, if it reaches the expected value of the quotation, which is calculated using the formula:
if FDOP > 0,

$$EV = O + (O * \mathrm{AMDU} * M)$$

if FDOP < 0,

$$EV = O + (O * \mathrm{AMDD} * M)$$

if FDOP equals 0, strategy does not open the position,
where: EV – expected value,
O – opening price,
AMDU – factor the average maximum deviation up price,
AMDD – factor the average maximum deviation down price,
M – multiplier, external parameter.

Second, if quotation does not achieve the desired level, position is closed at the time that is a parameter constrained in advance. External parameters: the minimum value of the FDOP to open position, the maximum value of the coefficient of VFDOP to open position, the multiplier.

Parameters in strategies allow to decide which neurons are worth using. The minimum value of the FDOP ratio (the minimum absolute value of the FDOP) allows to discard the neurons which historically have been characterized by high volatility. These neurons do not give an unambiguous signal to open positions in either direction. The maximum value of the VFDOP factor allows the rejection of neurons which at various analyzing periods have very diverged FDOP values. The signal from these neurons is also ambiguous, because in some periods gave unambiguous signals to open short positions at other times of the signals to open long positions. The last parameter is a multiplier. It allows to manipulate the expected value. And so if FDOP has too small value (that implicates small expected value → too early close → too little profit), we can increase the expected value using multiplier, so that positions are not closed too early (with too little profit). Multiplier for the strategy of the average deviation rate up/down works in a similar way. We can reduce or increase the value of the expected return of the transaction.

4. Experiments and discussions

The aim of the experiments was to review the effectiveness of the system. Network's neuron represents signals of opening and closing for 30 consecutive candles. Neural network consisted of 1024 neurons and was taught with a series of contracts FW20U09 by period of time from 18 June 2009 to September 18, 2009. Statistics of individual neurons in the network were taken from periods FW20Z7 (September 18, 2007 – December 18, 2007), FW20H8 (December 18, 2007 – March 18, 2008), FW20M8 (March 18, 2008 – June 18, 2008), FW20U8 (June 18, 2008 – September 18, 2008), FW20Z8 (September 18, 2008 – December 18, 2008), FW20H9 (December 18, 2008 – March 18, 2009), FW20M9 (March 18, 2009 – June 18, 2009), FW20U9 (June 18, 2009 – September 18, 2009). These were the periods of greatest activity of selected series of contracts. In other time periods, on each contracts series, volume was too low to take them into account.

The study was conducted with the maximum time to keep the position equal to 1 hour. The biggest profits gave neurons where the absolute value of the coefficient was enclosed in range <0.25, 0.3>, while the maximum deviation of the ratio enclosed in range <0.25, 0.55>. After defining the groups of neurons which had the highest profits, the author has studied factors that are responsible for closing the position for the last two strategies. The most profitable were transactions which had a expected value set to the percentage increase of the coefficient towards the opening position, multiplied by <2.1, 2.4>. However, for a strategy that closes the position at the average exchange rate variations in up/down the most effective factors were in the range of <1.6, 1.8>.

Table 1. Results of research (minimum FDOP = 0.25, the maximum VFDOP = 0.3; multiplier for Strategy 2 = 2.5; multiplier for Strategy 3 = 1.6)

Series	Reference price	Buy and hold	Percentage profit of "buy and hold"	Number of trans-actions	Maximum profit	Stra-tegy 1	Stra-tegy 2	Stra-tegy 3
FW20U9	1903	187	9.83	547	5508	9	–205.8	88.2
FW20M9	1460	446	30.55	683	6078	172	66.8	120.8
FW20H9	1822	–327	–17.95	1090	13282	705	–319.0	–272.0
FW20Z8	2458	–668	–27.18	625	10704	1632	761.0	1418.0
FW20U8	2690	–431	–16.02	376	4080	–146	28.6	2.6
FW20M8	2770	–52	–1.88	321	3768	–125	341.6	195.6
FW20H8	3501	–747	–21.34	588	10981	939	670.8	1189.8
FW20Z7	3613	–150	–4.15	268	3785	611	759.8	615.8
	Sum:	3008			58186	3797	2103.8	3358.8

Table 1 shows the results of the system for the chosen parameters (the minimum value of direction factor = 0.25, the maximum value of the variation direction factor = 0.3; multiplier factor for the strategy direction = 2.5; for Strategy deviation of the average rate up/down factor was 1.6). The table shows the results after deducting the cost of commissions.

A comparison of the results allows to conclude that the most profitable was the basic strategy which assumed keeping the position for one hour (Table 1 – Strategy 1,3797 profit points). "Buy and hold" strategy (Table 1 – "Buy and hold") for those periods, assuming correctly anticipate the direction of open positions, can earn 3008 profit points. The worst was the strategy of opening and closing positions by FDOP (Table 1 – Strategy 2,2103.8 profit points). As shown in Table 1, the basic advantage of the strategy was based on only one period FW20H9. During this period there was a pair of signals which generated much higher profits than expected. Strategies that predict the expected profit and close positions after reaching it did not allowed to gain full profits during this period. Although, in other studied periods closing position after gain expected profit, gave better results than those forced by the closing of a predetermined time interval.

5. Conclusion

The results obtained by the system encourage further research on the use of systems based on neural networks, in supporting investment decisions. Although the solution gave results slightly better than the strategy of "buy and hold" (direction has to be known prior to taking transaction) and in certain periods, brought a small loss, the author believes that the chosen research direction is promising. Strategy of "buy and hold" takes profits from the price changes between sessions, and presented solution assumed gaining the profits only by the price changes during the session.

Due to the fact that the actual signals delivered to the network are not scattered throughout the space of possible solutions, but are grouped in small, dense part of this space, it was difficult to cast on a two-dimensional space of SOM. So it is more reasonable to use neural gas or growing neural gas. Further research could focus on changing the form of the neuron, so that the last candles in defining string have greater importance – it would increase the sensitivity to what happened just prior to the signal, at the expense of what was happening a little earlier (it was more distant).

Another conclusion of the study is based on the facts that research on different networks learned well at the same period, gave different results and that many neurons gave ambiguous signals. Sequences of candle identified by the network as very similar (from the same cluster) gave signals that once prelude a significant rise in price and another time a significant fall. Slight impact on the learning of the network in the direction of associating patterns by similarity and profits which they gain can enhance the effectiveness of the characterized solution. Similarity of patterns was evaluated only on the Euclidean distance between them. The network did not focus on details that could be an unambiguous signal to open long/short position. One can conclude that the capture of such specific rules for specific signals can be identified by genetic algorithms, or gene expression programming.

Acknowledgements: I would like to express the deepest appreciation to professor Jerzy Korczak, whose encouragement, guidance and support enabled me to write the article. His knowledge and experience in the research area, artificial intelligence, and stock markets was invaluable.

References

Afolabi M.O., Olude O. (2007), *Predicting stock prices using a hybrid Kohonen Self Organizing Map (SOM)*, [in:] *Proceedings of the 40th Annual Hawaii International Conference on System Sciences*, IEEE Computer Society, Washington, pp. 814-821.

Bac M., Kwaśnicka H. (2009), Możliwości zastosowania algorytmów genetycznych w systemach informacyjnych wspomagających proces podejmowania decyzji gracza giełdowego, [in:] *Inżynieria i systemy ekspertowe*, Eds. A. Grzech, K. Juszczyszyn, H. Kwaśnicka, Akademicka Oficyna Wydawnicza Exit, Warszawa, pp. 663-676.

Ikeda Y., Tokinaga S. (2007), Multi-fractality analysis of time series in artificial stock market generated by multi-agent systems based on the genetic programming and its applications, [in:] *IEICE Transactions on Fundamentals of Electronics, Communications and Computer Sciences*, Oxford University Press, Oxford, pp. 2212-2222.

Khan A.U., Bandopadhyaya T.K., Sharm S. (2009), Classification of stocks using *Self Organizing* Map, [in:] *International Journal of Soft Computing Applications*, Issue 4, EuroJournals Publishing, pp. 19-24.

Korczak J., Lipiński P. (2006), Technology of intelligent agents used in financial data analysis, [in:] *Ogólnopolska Konferencja Naukowa Nowoczesne Technologie Informacyjne w Zarządzaniu*, NTIZ2006, Wydawnictwo Akademii Ekonomicznej, Wrocław, pp. 350-359.

Korczak J., Lipiński P. (2008), Systemy agentowe we wspomaganiu decyzji na rynku papierów wartościowych, [in:] *Rozwój informatycznych systemów wieloagentowych*, Ed. S. Stanek, Placet, Warszawa, pp. 289-301.

Li J., Tsang E.P.K. (1999), Improving technical analysis predictions: An application of genetic programming, [in:] *Proceedings of the Twelfth International Florida Artificial Intelligence Research Society Conference*, AAAI Press, USA, pp. 108-112.

Ritchie J.C. (1977), *Analiza fundamentalna*, Wydawnictwo Wig-Press, Warszawa.

Simunic K. (2003), *Visualization of stock market charts*, [in:] *Proceedings from The 11th International Conference in Central Europe on Computer Graphics, Visualization and Computer Vision*.

Stankevicius G. (2001), Forming of the investment portfolio using the self-organizing maps (SOM), *Informatica*, Vol. 12, No. 4, Vilnius, pp. 573-584.

WYKORZYSTANIE SIECI (SOM) DO WSPOMAGANIA KRÓTKOTERMINOWYCH DECYZJI INWESTYCYJNYCH

Streszczenie: Zainteresowanie rynkami giełdowymi na świecie wzrasta z roku na rok. Przyciągani możliwością szybkiego i łatwego zysku inwestorzy zwiększają płynność zarówno instrumentów bazowych, jak i instrumentów pochodnych, co budzi poczucie bezpieczeństwa i jeszcze bardziej zwiększa atrakcyjność rynków. Coraz bardziej zaawansowane techniki są wykorzystywane do wsparcia analizy instrumentów finansowych. Poza wykorzystywaniem komputerów do przeprowadzania standardowych analiz coraz częściej wykorzystuje się sztuczną inteligencję jako narzędzie do predykcji. W obliczu powyższych spostrzeżeń celami niniejszego artykułu są prezentacja i weryfikacja koncepcji zastosowania sieci neuronowych SOM do wspomagania decyzji inwestycyjnych.