1-1-2005

Machine learning techniques and use of event information for stock market prediction: A survey and evaluation

Paul D. Yoo
University of Technology, Sydney

Maria H. Kim
University of Wollongong, mhykim@uow.edu.au

Tony Jan
University of Technology, Sydney

Abstract

This paper surveys machine learning techniques for stock market prediction. The prediction of stock markets is regarded as a challenging task of financial time series prediction. In this paper, we present recent developments in stock market prediction models, and discuss their advantages and disadvantages. In addition, we investigate various global events and their issues on predicting stock markets. From this survey, we found that incorporating event information with prediction model plays very important roles for more accurate prediction. Hence, an accurate event weighting method and a stable automated event extraction system are required to provide better performance in financial time series prediction.

Keywords

survey, prediction, market, stock, information, event, evaluation, learning, techniques, machine

Disciplines

Business | Social and Behavioral Sciences

Publication Details

Yoo, P. D., Kim, M. H. & Jan, T. (2005). Machine learning techniques and use of event information for stock market prediction: A survey and evaluation. International Conference on Computational Intelligence for Modeling, Control and Automation (CIMCA 2005) (pp. 835-841). Piscataway, NJ: IEEE.

Machine Learning Techniques and Use of Event Information for Stock Market Prediction: A Survey and Evaluation

Paul D. Yoo, Maria H. Kim, Tony Jan
Department of Computer Systems
Faculty of Information Technology
University of Technology, Sydney
PO Box 123, Broadway, NSW 2007, Australia
dhyoo@it.uts.edu.au

Abstract

This paper surveys machine learning techniques for stock market prediction. The prediction of stock markets is regarded as a challenging task of financial time series prediction. In this paper, we present recent developments in stock market prediction models, and discuss their advantages and disadvantages. In addition, we investigate various global events and their issues on predicting stock markets. From this survey, we found that incorporating event information with prediction model plays very important roles for more accurate prediction. Hence, an accurate event weighting method and a stable automated event extraction system are required to provide better performance in financial time series prediction.

1. Introduction

Stock market prediction has been an important issue in the field of finance, engineering and mathematics due to its potential financial gain. As a vast amount of capital is traded through the stock market, the stock market is seen as a peak investment outlet. In addition, stock market prediction brings with it the challenge of proving whether the financial market is predictable or not. Since there has been no consensus on the validity of Efficient Market Hypothesis (EMH) which states the market is efficient and there is no space for prediction, researchers have strived for proving the predictability of the financial market [23].

With the advent of faster computers and vast information over the Internet, stock markets have become more accessible to either strategic investors or the general public. As the Internet provides a primary source of event information which has a significant impact on stock markets, the techniques to extract and use information to support decision making have become a critical task. To predict the stock market accurately, various prediction algorithms and models have been proposed by many researchers in both academics and industry. In this paper, recent development in prediction algorithms and models will be introduced and their performance will be compared. In addition, for accurate stock market prediction, we investigate various global events and their issues on predicting stock markets.

1.1. Background

Since the stock market was firstly introduced, many have attempted to predict the stock markets using various computational tools such as Linear Regression (LR), Neural Networks (NNs), Genetic Algorithms (GAs), Support Vector Machines (SVMs), Case-based Reasoning (CR) and others. Over the last decade, NNs have been most widely used and shown better performance over other approaches in many cases.

The earliest stock market prediction model based on NNs was implemented by White [45]. He used Feed Forward Neural Networks (FFNNs) to decode previously undetected regularities in the asset price movements such as fluctuations of common stock prices and showed how to search for such regularities using FFNNs. Since the initiative attempt by White, a number of researchers have participated in developing an accurate stock market prediction model. Phua et al. [31] used NNs with GA to predict the Singapore Stock Exchange Index and achieved accuracy rate of 81%. Kim and Han [18] also combined NNs with GA and predicted Korea Composite Stock Price Index 200. He achieved 82% of accuracy in predicting both weekly rising and declining stock market tendencies. GA has been investigated and shown to be effective in exploring a complex space in an adaptive way, guided by the evolution mechanism of reproduction, crossover,

Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Confe
Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'05)
0-7695-2504-0/05 $20.00 © 2005 IEEE

and mutation [9]. Recently, many researchers employed GA to improve the performance of AI techniques. For NNs, GA is popularly used to select NNs topology such as optimizing relevant feature subset, determining the optimal number of hidden layers and processing elements.

Many researchers have compared the NNs to the statistical approaches for pattern recognition [8] [35] [48]. Yoon et al. [47] pointed out the superiority of NNs over classical discriminant analysis. Garliauskas [8] also investigated stock market prediction using NNs in corporation with kernel function approach and the recursive prediction error method. He concluded that in predicting financial time series, NNs have better performance than classical statistical methods. In general, numerous studies have shown that NNs have the ability to predict stock markets more accurately than other methods [14] [30] [37] [48].

Although the former studies were confined to quantitative analysis, several recent studies have been performed based on the qualitative analysis. Many believed the integration of event-knowledge and NNs hold great promise for improved prediction in stock markets [7] [12] [13] [16] [19] [20] [27] [32]. Kohara [20] investigated ways of improving multivariate predictive models in stock price prediction using prior-knowledge and NNs. He used information from the newspaper headlines to improve the prediction ability. Their experimental results showed that the use of event-knowledge based on the prior-knowledge with NNs significantly reduced the prediction error rate on the 5% level of significance with profit by 40%.

In addition, Hong and Han [12] compared NNs with the event information to Random Walk (RW) model and NNs without event information. NNs with the event information have a figure of 0.527% in average error, which is a smaller error in comparison to other models. They proved that the NNs based forecaster is greatly superior to RW and that the effect of event information does exist. Furthermore, in predicting stock market, many researchers have recognized that qualitative factors such as political effects and international events played a very important role. Thus, they proposed that a stock market prediction system incorporate both quantitative and qualitative factors. Finally, their studies showed that NNs based on both quantitative and qualitative factors are far superior to the ones based only on the quantitative factors [21] [22].

1.2. Key issues in financial prediction

A number of researchers have shown their official viewpoint of EMH in academia [24] [45]. EMH states that the current market price reflects the assimilation of all the information available [10]. The information relevant to a market is contained in the prices and each time that new information arises, the market corrects itself and absorbs it, thus, the market is efficient and there is no space for prediction [5]. There has been a long ongoing debate about the validity of EMH, but no consensus has been made. In recent years, many researchers have claimed that EMH surely must be false [6]. Several studies have been performed on the data of stock markets in order to prove that the market is predictable. If any computational based systems have the capability to show reasonable prediction accuracy based on the market historical data, the validity of EMH can be questioned. Tsibouris and Zeidenberg [42] used NNs to predict stock market only based on past stock prices as inputs. Their empirical results showed some level of predictive ability and rejected the weak form of EMH. In addition, Eyden [3] developed a system which models the performance of the Johannesburg Stock Exchange and their system provided significant evidence by showing its capability to predict stock market directions, thus, it again dissented from EMH.

On the other hand, various studies were based on EMH. Fung et al. [7] investigated the immediate impact of news articles on the time series based on EMH. They presented an EMH based system to predict the future behavior of the stock market using non-quantifiable information (news articles) and the predictions were made only according to the contents of the news article. Finally, their empirical result indicated that the approach built based on EMH is profitable than simple trading strategy based on Buy-and-Hold. As it has been discussed, many of different studies have concluded to accept or refute EMH. Thus, the issue of market efficiency is still not fully investigated and it leads to a need for further research.

A number of researchers have used historical numeric time series data to predict stock markets and they achieved reasonable prediction accuracy while denying EMH [3] [42]. However, there are various factors that influence stock prices such as company's performance and robustness, trends of the market, investors' psychology, government involvement, changes in economic activity and so forth. Thus, many researchers have agreed to the existence of significant correlation between the events which represent above factors, and stock markets [7] [20] [27]. They have used the event information in addition to the numeric time series data and achieved some level of prediction accuracy. However, as each event has a different level of impact in terms of a time frame that the event can influence upon, each event should be weighted according to its level of impact which is defined based on previously built prior-knowledge. In addition, the web is seen as the primary event source for stock market prediction containing the latest and valuable event information. Thus, automated web mining techniques are required in order to yield higher prediction accuracy within short time frame.

2. Prediction methods

2.1. Traditional time series prediction

Traditional statistical models are widely used in economics for time series prediction. These models are capable of modeling linear relationships between factors that influence the market and the value of the market. In economics, there are two basic types of time series forecasting: univariate (simple regression) and multivariate (multivariate regression) [25].

One widely used example of univariate model is Box-Jenkins, which contains only one variable in the recurrent equation. It shows complicated process of fitting data to appropriate model parameters. In their equation, past values of moving average prices are included. A moving average model of order q is:

$$X_t = a_t + \theta_1 a_{t-1} + \theta_2 a_{t-2} + \dots + \theta_q a_{t-q}$$

Although Box-Jenkins shows good ability for short-term forecasting, it requires a large amount data to yield high accuracy [23]. Multivariate models are univariate models expanded to ‘discover causal factors that affect the behavior of the data’ [34]. As the name implies, the equation of the model contains more than one variable. A number of studies have compared the multivariate statistical model with NNs. Although the multivariate models had been widely used for predicting stock markets, several machine learning techniques are now replacing their roles.

In the literature, many researchers claimed that NNs substantially outperform traditional statistical methods [8] [35]. Lawrence [23] compared NNs with statistical and regression techniques. He used the JSE-system which uses NNs with GA. The JSE-system showed its capability for predicting market movement correctly 92% of the time, while Box-Jenkins only performed at a 60% accuracy rate. In addition, several systems using NNs showed consistently better performance than multiple linear regression models [33] [36]. One of the distinctive researches was performed by Yoon et al. [47]. Their NNs based system predicted the trend of stock price with 91% of accuracy as compared to 74% using multiple discriminant analysis (MDA). Thus, according to them, it can be concluded that NNs consistently outperform statistical and regression techniques.

2.2. Neural Networks

In the literature, it has been shown that NNs offer the ability to predict market directions more accurately than other existing techniques. The ability of NNs to discover non-linear relationships between the training input/output pairs makes them ideal for modeling non-linear dynamic systems such as stock markets [23].

One of the advantages is the ability to learn relationship through the data itself rather than assuming the functional form of the relationship. As NNs are known as a universal approximator, any relation can be modeled to any degree of accuracy when sufficient data for the modeling are given. In addition, it provides a level of tolerance to noisy and incomplete data representation. Another advantage is that NNs have non-linear, non-parametric adaptive learning properties and they have the most practical effect in modeling and forecasting. The non-linear nature of NNs shows great potential to solve many complex problems. Due to the above characteristics of stock markets, NNs can be applied to stock market prediction. Firstly, stock data is hard to model due to its complexity, thus, non-linear model is beneficial. Secondly, a large set of interacting input series is often required to explain a specific stock.

Regarding downsides, NNs have the black box problem, which does not reveal the significance of each variable and the way they weigh independent variables [23]. As the individual role of each variable cannot be determined, it is impossible to understand how the network produces future price of stock.

Another major problem with NNs is the overtraining problem. When NNs fit the data too well, the system loses the ability to generalize. Since the generalization ability of NNs is fundamental to predict future stock prices, overtraining is a serious problem. The overtraining usually occurs by two main reasons as NNs either have too many nodes or have too long training time period (epochs). However, overtraining can be prevented by performing test and train procedures or cross validation.

As for the tremendous noise and non-stationary characteristics in stock market data, Lawrence et al. [23] pointed out when the training of NNs is difficult for high noisy data, the networks fall into a naïve solution such as always predicting the most common output. Moreover, NNs have some limitations in

Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Confe
Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC’05)
0-7695-2504-0/05 $20.00 © 2005 IEEE

learning the patterns when input data have high dimensionality. Dash and Liu [2] put the emphasis on the feature selection and suggested that reducing the number of input variables sometimes lead to improved model performance for a given data set. The reduction and transformation of the irrelevant or redundant features may shorten the running time and yield more generalized results [2].

Recent researches tend to hybridize several AI techniques. Hiemstra [11] proposed fuzzy expert systems to predict stock market returns. He suggested that combining NNs and fuzzy logic capture the complexities of functional mapping and do not require the specification of the function to approximate. Tsaih et al. [41] integrated the rule based technique and NNs to predict the direction of change of the S&P 500 stock index futures on a daily basis.

2.3. Support Vector Machine

Support Vector Machine (SVM) based on the statistical learning theory, was developed by Vapnik [44] and his colleagues in the late 1970s. It has become a hot topic of intensive study due to its successful application in classification and regression tasks, especially in time series prediction and financial related applications [46].

SVM is a very specific type of learning algorithms characterized by the capacity control of the decision function, the use of the kernel functions and the sparsity of the solution [43]. Established on the unique theory of the structural risk minimization principle to estimate a function by minimizing an upper bound of the generalization error, SVM is shown to be very resistant to the overtraining problem, eventually achieving a high generalization performance.

Another key property of SVM is that training SVM is equivalent to solving a linearly constrained quadratic programming problem. Thus, the solution of SVM is relatively unique and globally optimal, unlike NNs training which requires nonlinear optimization with the danger of getting stuck at local minima.

Recently, several applications of SVM to financial forecasting problems have been reported [15] [38] [39]. One of the well known studies using SVM in stock market prediction was performed by Kim [15]. He applied SVM to financial forecasting and compared it with the backpropagation NNs and Case Based Reasoning (CBR). The experimental results showed that SVM outperformed backpropagation NNs and CBR. As SVM implements the structural risk minimization principle, SVM leads to better generalization than traditional techniques.

2.4. Case Based Reasoning

Although NNs offer relatively good learning ability than other techniques, they cannot always explain why they arrive at a particular solution. Moreover, they cannot always guarantee a completely certain solution, arriving at the same solution repeatedly given the same training data, nor can guarantee the best solution [40]. However, unlike NNs, expert systems typically provide explanations for their solutions. Expert systems primarily capture the knowledge of individual experts. Organizations have collective knowledge and expertise they have built up over the years. This knowledge can be captured and stored using CBR. CBR is a reasoning technique that reuses past cases to find a solution to the new problem. CBR not only captures organization knowledge but also provides explanations for the derived solutions. For this reason, CBR is popularly applied to many applications.

Kim [17] proposed a new hybrid model of GA and CBR for stock market prediction. From his preliminary studies, he found feature weighting or feature subset selection are very important to enhance the prediction performance of the CBR system. Thus, they used GA as a method of feature subset selection in the CBR system. They compared the results of four different models to test the effectiveness of the proposed model. Compared models are conventional CBR (COCBR), feature weighting using GA for CBR (FWCBR), feature selection using GA for CBR (FSCBR) and simultaneous optimization using GA for CBR (SOCBR). The empirical results showed that SOCBR achieves higher prediction accuracy than COCBR, FWCBR and FSCBR. He concluded that the hybrid model of GA and CBR offers a viable alternative approach to stock market prediction.

3. Event information

In stock markets, there are many factors that can influence the share price. These factors can be derived from the news release about small companies or the news of superpower national economy. These incidents are called ‘events’ [29]. The primary reason of incorporating event knowledge in stock market prediction is based on an assumption that the future price of a stock partially depends on various political and international events as alongside the various economic indicators. Thus, many studies have used event information (qualitative factors) as well as quantitative data in predicting stock markets.

One of the popular studies using the prior-knowledge and event-knowledge was performed by Kohara et al. [20]. They incorporated prior-knowledge in stock prediction such as newspaper information on domestic and foreign events. Event-knowledge is extracted from the news paper headlines in accordance with certain prior-knowledge (See Fig. 1). Prior-knowledge is the information that stems from previous experience. Thus, based on the prior-knowledge, decisions can be made whether a particular event can positively influence the stock market tendencies or not.

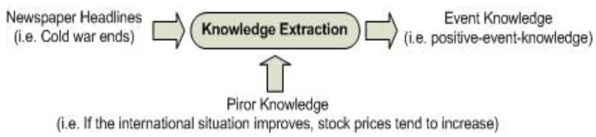

Fig. 1. Extracting Event Knowledge

Kohara et al. [20] selected several economic indicators (interest rate, price of crude oil, and New York Dow Jones average of the closing price) and fed them together with event-knowledge into NNs. Their experimental results showed incorporation of event-knowledge improved the prediction ability of NNs by reducing the error rate on the 5% level of significance.

Meanwhile, how to incorporate the impact from news information into time series models is crucial. Maheu and McCurdy [26] specified a GARCH Jump model for return series, which can be directly measured from price data. The latent news process is postulated to have two separate components, normal and unusual news events. These news innovations are identified through their impact on return volatility. The unobservable normal news innovations are assumed to be captured by the return innovation component, $\varepsilon_{1,t}$. This component of the news process smoothly causes evolving changes in the conditional variance of returns. The second component causes infrequent large moves in returns, $\varepsilon_{2,t}$. The impacts of this unusual news are labeled as jumps. Given an information set at time *t-1*, consisting of the history of returns $\Phi_{t-1} = \{r_{t-1}, \ldots, r_t\}$, the two stochastic innovations, $\varepsilon_{1,t}$ and $\varepsilon_{2,t}$ drive returns $r_t = \mu + \varepsilon_{1,t} + \varepsilon_{2,t}$. $\varepsilon_{1,t}$ is a mean-zero innovation $E[\varepsilon_{1,t} \mid \Phi_{t-1}] = 0$ with a normal stochastic forcing process, $\varepsilon_{1,t} = \sigma_t z_t, z_t \sim NID(0,1)$ and $\varepsilon_{2,t}$ is a jump innovation.

The previous models provide general frameworks to incorporate the impacts from news articles. But these methods do not provide an approach to figuring out the influential or significant news of a given stocks in the face of thousands of news articles from all kinds of resources. Therefore, these methods cannot make significant improvement in practice.

In the literature, a number of researchers stated that stock prices are significantly correlated with the event information and many attempted to use both the event information and numeric time series data as input data. Fawcett and Provost [4] formulated the stock forecasting problem as an activity monitoring the relationship between the news articles and stock prices. Fung et al. [7] proposed a system that predicts the changes of stock trends by analyzing the influence of qualitative information (news articles). In particular, they investigated the immediate impact of news articles on the time series based on the EMH. They proposed a new statistical based piecewise segmentation algorithm to identify trends on the time series. In the literature, there is a significantly growing recognition of incorporating prior and event-knowledge into prediction models such as NNs.

4. Event information on the Web

The evolution of the Internet with the global information infrastructure has led to an explosion in the amount of available information. Enormous event information which may have great influence on stock markets is available on the web. Whereas newspapers are updated once or twice a day, the real time news sources are frequently updated on the spot. This information not only contains global and regional news but also valuable citations from influential bankers, politicians and financial analysts. The information which consists of the news and the citations are seen as primary movers of bond, stock and currency market over the world. Hong and Han [13] stated that as the popularity of the Internet increases, many newspapers expand their services by providing news information on the web in order to be more competitive and increase profit. News information includes articles on the political situation, social conditions, international events, government policies, trader's psychology, and so forth, which we see and understand through the Internet. Such information is formulated in the form of texts, referred to as documents, thus, text mining is required.

Hong and Han [13] introduced an automated system (KBNMiner) that acquires event-knowledge from the Internet for the prediction of interest rates. This study shows a clear concept of event-knowledge using an automated system. The KBNMiner is designed to adopt a prior-knowledge base, which is seen as 'expert knowledge' as a foundation on which to collect the event information from the Internet automatically and then to apply the information to a neural network model for interest rate prediction. What is notable from

Proceedings of the 2005 International Conference on Computational Intelligence for Modelling, Control and Automation, and International Confe
Intelligent Agents, Web Technologies and Internet Commerce (CIMCA-IAWTIC'05)
0-7695-2504-0/05 $20.00 © 2005 IEEE

this study is that the web mining technique they applied to predicting interest rates can also be used for stock market prediction.

5. Conclusion and future work

In this paper, we examined recent developments in stock market prediction models. By comparing various prediction models, we found that NNs offer the ability to predict market directions more accurately than other existing techniques. The ability of NNs to learn non-linear relationships from the training input/output pairs enables them to model non-linear dynamic systems such as stock markets more precisely [23].

Other models such as SVM and CBR have also become popular in stock market prediction. SVM showed its successful application in classification task and regression tasks, especially on time series prediction and financial related applications [46]. CBR is a reasoning technique that reuses past cases to find a solution to the new problem. CBR captures organization knowledge and expertise while providing explanations for the derived solutions. For this reason, CBR is popularly applied to many applications.

In addition, by studying several important issues in stock markets, we found that that many researchers have recognized that qualitative factors such as political effects and international events can have a significant impact on stock prices. It has been reviewed that NNs based on both quantitative and qualitative factors are far superior to the ones based only on the quantitative factors. In addition, the web is regarded as the primary event source for stock market prediction containing the latest and latent event information. Thus, for the stock market prediction, a level of web mining technique is required in order to yield higher prediction accuracy and to make prediction in short time frame.

For further research, firstly, prior-knowledge database should be built by analyzing historical events on stock markets. Based on the prior-knowledge, the event weighting schema will be developed and each event should be weighted accordingly. Finally, a proposed model which incorporates the weighed events into the numeric time series data should be compared empirically with other models.

6. References

[1] Baestaens, D. E. and Bergh, W. 1995, ‘Tracking the Amsterdam stock index using neural networks’, Neural Networks in the Capital Markets, vol 10, pp 149–162.

[2] Dash, M. and Liu, H. 1997, ‘Feature selection for classifications’, Intelligent Data Analysis: An International Journal, vol. 1, pp. 131-156.

[3] Eyden, R. J. 1996, ‘The Application of Neural Networks in the Forecasting of Share Prices’, Finance and Technology Publishing.

[4] Fawcett, T., and Provost, F.J. 1996. ‘Combining data mining and machine learning for effective user profiling’, In Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, KDD-96, pp. 8-13.

[5] Fama, E. 1970, ‘Efficient Capital Markets: A Review of Theory and Empirical Work’, Journal of Finance, vol. 25, pp. 383-417.

[6] Fama, E. F. 1991, ‘Efficient capital markets: II’. Journal of Finance. vol 47, pp. 1575-1617.

[7] Fung, G. P. C., Yu, J. X. and Lam, W. 2002, ‘News Sensitive Stock Trend Prediction’, Lecture Notes in Computer Science, vol. 2336, Jan 2002, pp. 481.

[8] Garliauskas, A. 1999, ‘Neural Network Chaos and Compuational Algorithm of Forecast in Finance’, Proceedings of the IEEE SMC Conference on Systems, Man, and Cybernetics 2, pp. 638-643, 12-15 October.

[9] Goldberg, D. E. 1989, ‘Genetic algorithm in search, optimization, and machine learning’, Reading, MA: Addison-Wesley.

[10] Hellstrom, T. and Holmstrom, K. 1998, ‘Predicting the Stock Market’, Technical Report Series IMa-TOM-1997-07.

[11] Hiemstra, Y. 1995, ‘Modeling structured nonlinear knowledge to predict stock market returns’, R. R.(ed.), Chaos & Nonlinear Dynamics in the Financial Markets: Theory, Evidence and Applications, Irwin, pp. 163-175.

[12] Hong, T. and Han, I. 2002, ‘Knowledge-based data mining of news information on the Internet using cognitive maps and neural networks’, Expert Systems with Applications, vol. 23, no. 1, pp. 1-8.

[13] Hong, T. and Han, I. 2004, ‘Integrated approach of cognitive maps and neural networks using qualitative information on the World Wide Web: KBN Miner’, Expert Systems, vol. 21 no.5, pp. 243-252.

[14] Kim, J. W. Weistroffer, H. R. and Redmond, R. T. 1993, ‘Expert Systems for Bond Rating: A Comparative Analysis of Statistical, Rule-Based, and Neural Network Systems’, Expert Systems, vol. 10, pp. 167-171.

[15] Kim, K. 2003, ‘Financial time series forecasting using support vector machines’, Neurocomputing, vol. 55, pp. 307-319.

[16] Kim, K. Hong, T. & Han, I. 1998, ‘Knowledge Discovery Process In Internet For Effective Knowledge Creation: Application To Stock Market’, Korea Advanced Institute of Science and Technology.

[17] Kim, K. 2004, ‘Toward Global Optimization of Case-Based Reasoning Systems for Financial Forecasting’, Applied Intelligence, vol. 21, no. 3, pp. 239-249.

[18] Kim, K. and Han, I. 2000, ‘Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index’, Expert System Appliance, vol. 19.

[19] Kohara, K. 2003, ‘Selective-Learning-Rate Approach for Stock Market Prediction by Simple Recurrent Neural

Networks', Lecture Notes in Computer Science, vol 2773, Jan 2003, pp. 141-147.

[20] Kohara, K. Ishikawa, T. Fukuhara, Y. and Nakamura, Y. 1997, 'Stock Price Prediction Using Prior Knowledge and Neural Networks', Intelligent System In Accounting, Finance and Management, vol. 6, pp. 11-22 vol.16.

[21] Kuo, R. J. Chen, C. H. and Hwang, Y. C. 2001, 'An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network', Fuzzy Sets and Systems, vol. 118, pp. 21-45.

[22] Kuo, R. J. Lee, L. C. and Lee, C. F. 1996, 'Integration of Artificial Neural Networks and Fuzzy Delphi for Stock Market Forecasting', IEEE, June, pp. 1073-1078.

[23] Lawrence, R. 1997, 'Using Neural Networks to Forecast Stock Market Prices', University of Manitoba.

[24] Lowe, D and Webb, A. R. 1991, 'Time series prediction by adaptive networks: A dynamical systems perspective. IEEE Computer Society Press.

[25] Maddala, G. S. 1992, 'Introduction to economics', Macmillan Publishing Company, New York, Toronto.

[26] Maheu, J. M. & McCurdy, T. H. 2004, 'News Arrival, Jump Dynamics and Volatility Components for Individual Stock returns', Journal of Fiannace, vol 2, pp. 755.

[27] Mittermayer, M. A. 2004, 'Forecasting intraday stock price trends with text mining techniques', System Sciences, 2004. Proceedings of the 37th Annual Hawaii International Conference on, pp. 64-73.

[28] Mulhern, F. J. and Caprara, R. J. 1994, 'A nearest neighbor model for forecasting market response', International Journal of Forecasting, vol. 10, pp. 191-207.

[29] Ng, A. and Fu, A.W. 2003, 'Mining Frequent Episodes for Relating Financial Events and Stock Trends', Lecture Notes in Computer Science, vol. 2637, pp. 27-39.

[30] Patuwo, E. Hu, M. Y. and Hung, M. S. 1993, 'Two-Group Classification Using Neural Networks', Decision Sciences, vol. 24, pp. 825-845.

[31] Phua, P. K. H. Ming, D. and Lin, W. 2000, 'Neural Network with Genetic Algorithms for Stocks Prediction', Fifth Conference of the Association of Asian-Pacific Operations Research Societies, 5th - 7th July, Singapore.

[32] Pui Cheong Fung, G. Xu Yu, J. and Lam, W. 2003, 'Stock prediction: Integrating text mining approach using real-time news', Computational Intelligence for Financial Engineering, 2003. Proceedings. 2003 IEEE International Conference on, pp. 395-402.

[33] Refenes, A. Zapranis, A.D. and Francis, G. 1995. 'Modelling stock returns in the framework of APT: A comparative study with regression models'. Neural Networks in the Capital Markets, vol. 7, pp. 101–126.

[34] Robert, J and Van, E. 1996, 'The Application of Neural Networks in the Forecasting of Share Prices', Finance and Technology Publishing.

[35] Schumann, M. and Lohrbach, T. 1993, 'Comparing artificial neural networks with statistical methods within the field of stock market prediction', System Sciences, 1993, Proceeding of the Twenty-Sixth Hawaii International Conference on, vol. 4, pp. 597-606 vol.594.

[36] Steiner, M and Wittkemper, H. 1995, 'Neural networks as an alternative stock market model', Neural Networks in the Capital Markets, vol. 9, pp. 137–148.

[37] Subramanian, V. Hung, M. S. and Hu, M. Y. 1993, 'An Experimental Evaluation of Neural Networks for Classification', Computers and Operations Research, vol. 20, pp. 769-782.

[38] Tay, F. and Cao, L. J. 2001, 'Application of support vector machines in financial time series forecasting'. Omega, vol. 29, pp. 309-17.

[39] Tay, F, and Cao, L. 2001, 'A comparative study of saliency analysis and genetic algorithm for feature selection in support vector machines'. Intelligent Data Analysis, vol. 5, pp. 191-209.

[40] Trippi, R. R. and Turban, E. 1992, 'Neural Network in Finance and Investing', Probus.

[41] Tsaih, R. Hsu, Y. & Lai, C. C. 1998, 'Forecasting S&P 500 stock index futures with a hybrid AI system', Decision Support Systems, vol. 23, pp. 161-174.

[42] Tsibouris, G and Zeidenberg, M. 1995 'Testing the Efficient Markets Hypothesis with gradient descentalgorithms'. In Neural Networks in the Capital Markets, vol. 8, pp 127–136.

[43] Vapnik, V. 1998, 'Statistical learning theory', Wiley, New York.

[44] Vapnik, V. 1999, 'An overview of statistical learning theory'. IEEE Transactions of Neural Networks, vol. 10, pp. 988-99.

[45] White, H. 1988, 'Economic prediction using neural networks: a case of IBM daily stock returns', IEEE International Conference on Neural Networks, vol. 2, pp. 451-458.

[46] Yang, H. Chan, L. and King, I. 2002, 'Support Vector Machine Regression for Volatile Stock Market Prediction', Lecture Notes in Computer Science, vol. 2412, pp. 391.

[47] Yoon, Y. Swales, G. Jr. and Margavio, T.M. 1993, 'A Comparison of Discriminant Analysis Versus Artificial Neural Networks', Journal of the Operational Research Society, vol. 44, pp. 51-60.

[48] Yoon, Y. and Swales, G. 1991, 'Predicting Stock Price Performance: A Neural Networks Approach', Proceedings of the IEEE Twenty-Fourth Annual Hawaii International Conference on System Sciences, vol. 4, pp. 156-162.

0-7695-2504-0/05 $20.00 © 2005 IEEE