Predicting the Australian Stock Market Index Using Neural Networks Exploiting Dynamical Swings and Intermarket Influences

Heping Pan, Chandima Tilakaratne, and John Yearwood

School of Information Technology & Mathematical Sciences,
University of Ballarat
Mt Helen, Victoria 3350, Australia
h.pan@ballarat.edu.au

Abstract. This paper presents a computational approach for predicting the Australian stock market index - AORD using multi-layer feed-forward neural networks from the time series data of AORD and various interrelated markets. This effort aims to discover an optimal neural network or a set of adaptive neural networks for this prediction purpose, which can exploit or model various dynamical swings and inter-market influences discovered from professional technical analysis and quantitative analysis. Four dimensions for optimality on data selection are considered: the optimal inputs from the target market (AORD) itself, the optimal set of interrelated markets, the optimal inputs from the optimal interrelated markets, and the optimal outputs. Two traditional dimensions of the neural network architecture are also considered: the optimal number of hidden layers, and the optimal number of hidden neurons for each hidden layer. Three important results were obtained: A 6-day cycle was discovered in the Australian stock market; the time signature used as additional inputs provides useful information; and a minimal neural network using 6 daily returns of AORD and 1 daily returns of SP500 plus the day of the week as inputs exhibits up to 80% directional prediction correctness.

1 Introduction

Predicting financial markets has been one of the biggest challenges to the AI community since about two decades ago. The objective of this prediction research has been largely beyond the capability of traditional AI because AI has mainly focused on developing intelligent systems which are supposed to emulate human intelligence. However, the majority of human traders cannot win consistently on the financial markets. In other words, human intelligence for predicting financial markets may well be inappropriate. Therefore, developing AI systems for this kind of prediction is not simply a matter of re-engineering human expert knowledge, but rather an iterative process of

T.D. Gedeon and L.C.C. Fung (Eds.): AI 2003, LNAI 2903, pp. 327-338, 2003.

knowledge discovery and system improvement through data mining, knowledge engineering, theoretical and data-driven modeling, as well as trial and error experimentation.

Multi-layer feed-forward neural networks, also known as multi-layer Perceptrons, are sub-symbolic connectionist models of AI, which have been proven both in theory and in practical applications to be able to capture general nonlinear mappings from an input space to an output space. The capacity of neural networks to represent mappings was investigated by Cybenko [3, 4] and Hornik et al. [6]. It is now known that a single nonlinear hidden layer is sufficient to approximate any continuous function, and two nonlinear hidden layers are enough to represent any function.

Using neural networks to predict financial markets has been an active research area since the 1990s [1-2, 5, 7-9, 11-26]. Most of these published works are targeted at US stock markets and other international financial markets. Very little is known or done on predicting the Australian stock market. From our empirical knowledge of technical and quantitative analysis and our first-hand observations in the international stock markets, it appears clearly to us that every stock market is different, and has its unique "personality" and unique position in the international economic systems. While the empirical knowledge gathered from predicting US and other international markets is definitely helpful, developing neural networks particularly devoted to predicting the Australian stock market requires highly specialized empirical knowledge about this market and its dynamic relationships to the US stock markets and other international financial markets.

There are quite a few non-trivial problems involved in developing neural networks for predicting the Australian stock market. Two traditional problems are the optimal number of hidden layers and the optimal number of hidden neurons for each hidden layer. Theoretically sound and practically sufficient solutions for these two problems can be found. However, the most critical problem in this research is the selection of the optimal input vector of features to be extracted from the time series data of the market. In our case, the primary source of data is the Australian stock market index, either the Australian all ordinary index (AORD), or the ASX/S&P 200 index. The secondary sources of data are US stock market indices and other international financial market indices. From the perspective of a trading system development, the second most important problem is the optimal prediction horizon, or the time frame. That is how far into the future the neural networks can predict with the highest reliability.

The remainder of this paper is organized as follows: Section 2 defines the problem of this prediction research and formalizes the major sub-problems; Section 3 describes the neural network approach to the prediction problem; Sections 4 to 5 present our solutions to the problems of optimal input and output selection; Section 6 points out further possibilities for research; Section 7 concludes the paper.

2 The Stock Index Prediction Problem

Let $X(t)$ be the target index at the current time *t*, note that $X(t)$ is a vector of five components:

$$X(t) = (X.O(t), X.H(t), X.L(t), X.C(t), X.V(t)) \tag{1}$$

where O, H, L, C, V denote respectively the open, high, low, close index level and the traded volume for the trading period of time *t*. In general, we take each single day as the standard resolution of time, so we use daily charts as the standard data. Let $Y_k(t)$ be the index of another market which is considered to be interrelated to the target market $X(t)$, where $k = 1,2,\cdots,K$, and K denotes the total number of interrelated markets selected. Similarly, $Y_k(t)$ is a vector:

$$Y_k(t) = (Y_k.O(t), Y_k.H(t), Y_k.L(t), Y_k.C(t), Y_k.V(t)) \tag{2}$$

We shall call $Y_k(t)$ an inter-market of $X(t)$, and

$$Y(t) = \{Y_k(t) \mid k = 1,2,\cdots K\} \tag{3}$$

the inter-markets – the complete set of selected inter-markets of $X(t)$. We assume the availability of historical time series data, usually called charts in technical analysis, $DX(t)$ of the target market and $DY(t)$ of the inter-markets, defined as

$$DX(t) = \{X(t) \mid t = t-N+1, t-N+2, \cdots, t-2, t-1, t\} \tag{4}$$

$$DY(t) = \{DY_k(t) \mid k = 1,2,.\cdots K\} \tag{5}$$

where

$$DY_k(t) = \{Y_k(t) \mid t = t-N+1, t-N+2, \cdots, t-2, t-1, t\} \tag{6}$$

and N is the total number of trading days for the standard time resolution (or minutes for intra-day charts). The time t starts from $N-1$ trading days back into the past, taking today as t. So the current time is just after today's market close for the target market and all the inter-markets.

The problem of prediction is to use the given historical chart data $DX(t)$ and $DY(t)$ of the last N trading days to predict the index of the target market T days into the future

$$(DX(t), DY(t)) \mapsto X(t+T) \tag{7}$$

where T can range from 1 to 10 for short-term prediction, or to 22 for monthly prediction, or to 256 for yearly prediction, and so on. Chaos theory tells us that the precision and reliability of prediction decays exponentially in time.

3 A Neural Network Approach for Stock Index Prediction

The approach of using neural networks for predicting the stock market index starts with defining a mapping M from a n-dimensional input space $\{x(x_1, x_2, \cdots, x_n)\}$ to an m-dimensional output space $\{z(z_1, z_2, \cdots, z_m)\}$:

$$M : x(x_1, x_2, \cdots, x_n) \mapsto z(z_1, z_2, \cdots, z_m) \tag{8}$$

Here we assume we have defined n real-valued features from the available data-set $DX(t)$ and $DY(t)$ at the current time t, and we want to predict m real values which may correspond to one or more target index levels in the future.

There are only two sensible architectures of neural network to consider for the stock index prediction problem: if we assume the stock index prediction problem can be modeled as a nonlinear continuous function (mapping), we should use a three-layer Perceptron with one hidden nonlinear layer, and this is the basic architecture; however, in general, this mapping should be assumed to be just any function which may contain discontinuities, so the general architecture should be a four-layer Perceptron with two hidden nonlinear layers. In both cases, the output layer should be linear because the outputs can be positive or negative real values. Pan and Foerstner [10] proposed an MDL-principled approach for determining the bounds of the number of hidden neurons.

The three-layer Perceptron with one hidden nonlinear layer of h hidden neurons can be represented as follows:

$$z_j = \sum_{k=1}^{h} w_{kj} y_k + b_j , \quad \text{for } j = 1,2,\cdots,m \tag{9}$$

$$y_k = \phi(\sum_{i=1}^{n} w_{ik} x_i + a_k), \text{ for } k = 1,2,\cdots,h \tag{10}$$

where y_k is the output of the k-th hidden neuron, a_k, b_j are the bias for the k-th hidden neuron and the j-th output neuron respectively, $\phi(\cdot)$ is the nonlinear transfer function for the hidden neurons, which generally takes the form of sigmoid function

$$\phi(x) = \frac{1}{1+e^{-x}} \tag{11}$$

For the four-layer Perceptron with two hidden nonlinear layers, we have similar formulas as (10) for each of the hidden layers, but the two hidden layers may have different numbers of hidden neurons.

4 Inspiration and Data Mining for Optimal Input Selection

The vector of input features includes two sub-vectors: those features extracted from the target market index and those extracted from the inter-market indices. The basic candidates for the features from the target market index are the relative returns of the closing index for the short term and geometrically spanning over the intermediate term. The relative return of the time interval τ is defined as

$$r_\tau(t) = \frac{X.C(t) - X.C(t-\tau)}{X.C(t-\tau)} \tag{12}$$

where $r_1(t)$ refers to daily returns, $r_5(t)$ to weekly returns, $r_{22}(t)$ to monthly returns, and so on. Our empirical knowledge from technical analysis suggests the following candidates:

$$(r_1(t), r_1(t-1), r_1(t-2), r_1(t-3), r_1(t-4), r_1(t-5)) \tag{13}$$

$$(r_5(t-6), r_5(t-11), r_5(t-16), r_5(t-21), r_5(t-26), r_5(t-31), \cdots)) \tag{14}$$

A more general approach for spanning over the past through a geometrical scale space is to use Fibonacci ratios, such as

$$(r_2(t-6), r_3(t-8), r_5(t-11), r_8(t-16), r_{13}(t-24), r_{21}(t-37), \cdots) \tag{15}$$

It is known that the relative return series of (13) and (14) do not have a long-term memory, so they may only capture short-term market cycles and candlestick chart patterns. One of the most significant weaknesses of using relative return series is the inability of representing important index support/resistance levels spanning over intermediate to longer terms. To overcome this weakness, we should consider the second form of relative return:

$$q_\tau(t) = \frac{X.C(t) - X.C(t-\tau)}{X.C(t)} \tag{16}$$

Then a vector similar to (13) can be defined as

$$(r_1(t), r_2(t), r_3(t), r_4(t), r_5(t), r_6(t))$$

Similarly features of geometrical time span can be defined.

According to our knowledge of inter-market technical analysis and our observations of the international stock markets and other financial markets, we have selected the following international markets as the inter-markets of the target market:

$$Y_1 = \text{US S\&P 500 Index} \tag{17}$$

$$Y_2 = \text{US Dow Jones Industrial Average Index} \tag{18}$$

$$Y_3 = \text{US NASDAQ 100 Index} \tag{19}$$

$$Y_4 = \text{Gold \& Silver Mining Index} \tag{20}$$

$$Y_5 = \text{AMEX Oil Index} \tag{21}$$

The space of possible inter-markets for the Australian stock market is not limited to these, however, these inter-markets of (17)-(21) provide the most influential parts of the global economical system affecting the Australian stock market. Candidates of the additional input features selected from these inter-markets are the relative returns in the form of either (12) or (16). If the prediction horizon is just the next day, the relative returns of the inter-markets at the latest time t are sufficient; if the horizon is longer, then at least a short-term time series of the relative returns of one of the inter-markets should be used, together with the relative return time series of the target market such as (13) or (15).

Figures 1 & 2 show the autocorrelation and partial autocorrelation of the target market X = Australian All Ordinary Index (AORD) relative return as defined by (12). From the autocorrelation and partial autocorrelation figures, we can clearly see that there is a 6-day cycle in the Australian stock market. This is surprising as we usually assume there is a 5-day cycle, known as the Day of the Week.

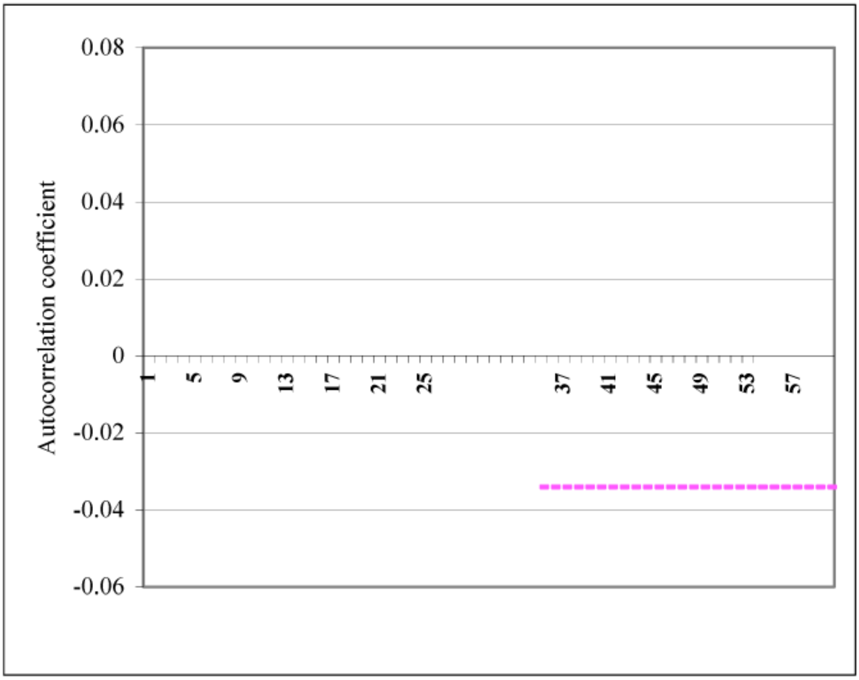

Fig. 1. Autocorrelation function of X

Figures 3 & 5 show the cross correlation between X and each Y_k, $k = 1,2,3$ as defined by (17)-(19). The correlations between $X(t)$ and $Y_k(t-1)$, for $k = 1,2,3$, are extremely high, and those between $X(t)$ and $Y_k(t)$ are also very significant. Significant correlations between $Y_1(t)$ and each of $Y_4(t), Y_5(t)$ are also found as expected according to our empirical knowledge. Therefore, $Y_4(t), Y_5(t)$ should also be included in the inter-markets of $X(t)$.

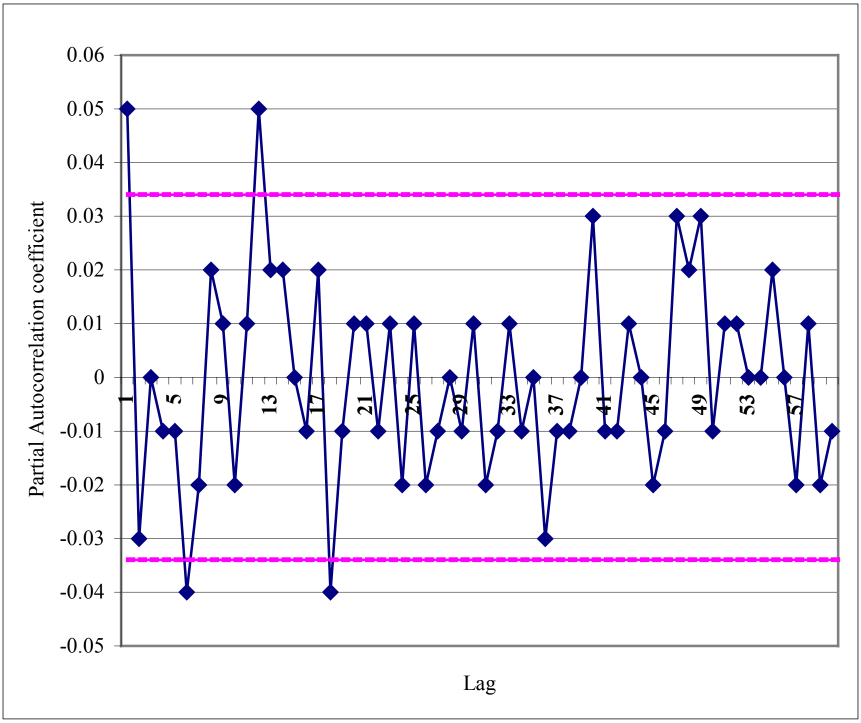

Fig. 2. Partial autocorrelation coefficients of X

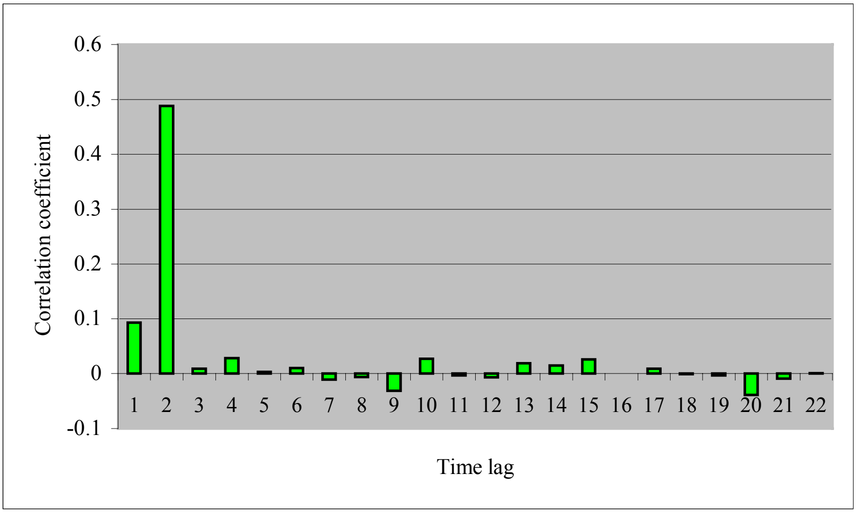

Fig. 3. Cross correlation between X and Y_1 = S&P 500

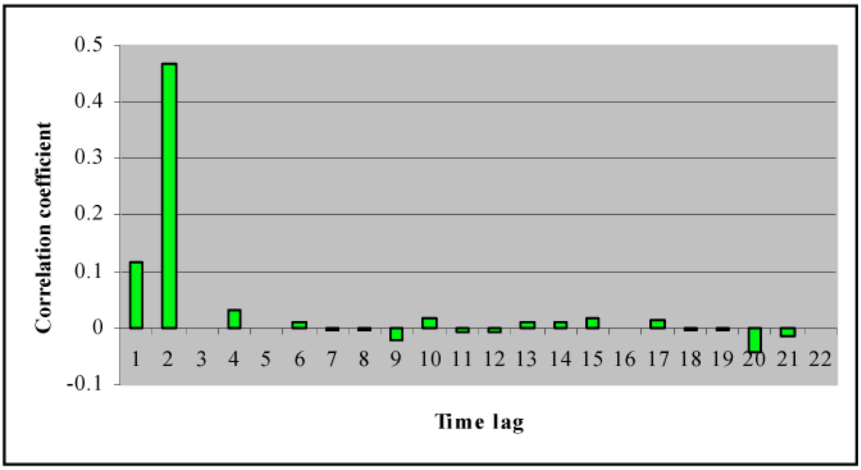

Fig. 4. Cross correlation between X and Y_2 = Dow Jones

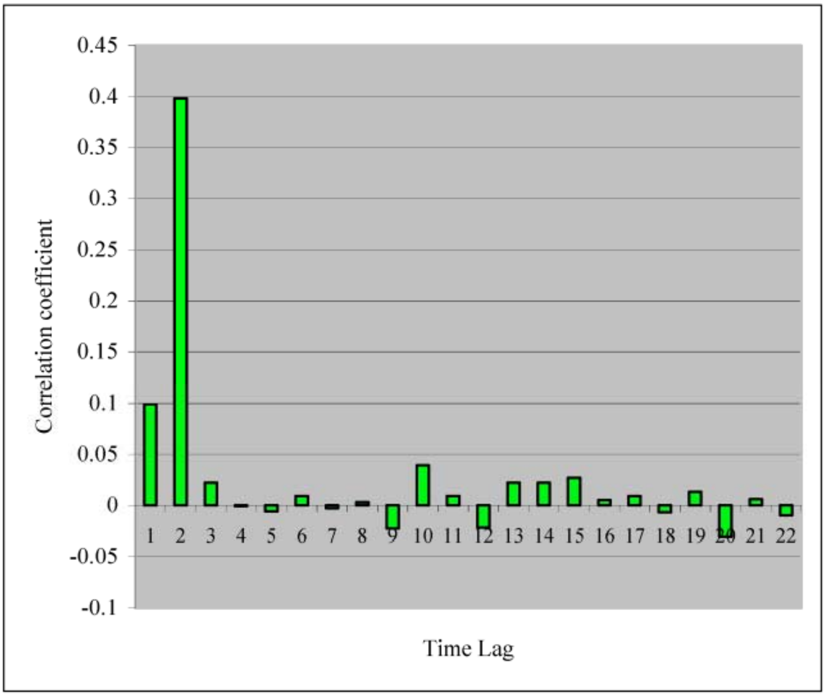

Fig. 5. Cross correlation between X and Y_3 = NASDAQ

5 Inspiration and Experimentation for Optimal Output Selection

With a series of short-term daily and weekly returns as input vector, we expect to be able to only predict 1 to 10 days into the future. Therefore, we have limited our output candidates to be $\{X(t+1), X(t+2), \cdots, X(t+10)\}$. According to the result of our statistical data mining (Figures. 1& 2), the Australian stock index, AORD, shows clearly cyclic behaviour with period 6 days. Therefore, we decided to use the last 6 daily returns of AORD in the input vector. Among the inter-markets (17)-(21) considered, Y_1 - (US S&P 500 Index) shows the highest correlation to the target market. So it was also included in the input vector. In addition, the day of the week as a time signature was also included as an additional input in our experiments. The number of hidden layers are limited to 1 or 2, and the number of hidden neurons for each hidden layer ranges from 2 to 50 in our experiments.

The historical index time series data of $DX(t)$ and $DY(t)$ between January 1990 and May 2003 were used. The whole data set is divided into two data sets: 20% of the data are randomly selected and put into the test data set, and the remaining 80% into the training data set.

The performance of neural networks are measured in terms of root mean square error (RMSE) and variance reduction (VR) for the training data set, and the sign correctness percentage (SCP) for testing the trained network on the test data set. Let N_1, N_2 be the number of data points for the training data and the test data respectively, then $N_1 + N_2 = N$ is the total number of data points. The RMSE and VR for the training data are calculated as

$$RMSE = \sqrt{\frac{1}{N_1}\sum_{k=1}^{N_1}(z_k - o_k)^2} \tag{22}$$

$$VR = \left(1 - \frac{\sum_{k=1}^{N_1}(z_k - o_k)^2}{\sum_{k=1}^{N_1}(z_k - \bar{z})^2}\right)\cdot 100\% \tag{23}$$

where z_k, o_k are the desired and actually calculated output for the k-th data point in the training data set, $\bar{z}$ is the mean of $\{z_k \mid k = 1,2,\cdots,N_1\}$. The SCP for the testing data set is defined as

$$SCP = \frac{|\{sign(z_k) = sign(o_k) \mid k = 1,2,\cdots,N_2\}|}{N_2} \tag{24}$$

where $|\{\}|$ denotes the number of elements in the given set.

A number of neural networks were trained. The best result is achieved on predicting the next day market direction and magnitude $X(t+1)$ with the selected inputs. However, it must be pointed out that prediction of different time frames requires different input selections, which may lead to significantly different network architectures. Table 1 shows the results of using different inputs with 1 hidden layer of 2 hidden neurons for predicting $X(t+1)$:

Table 1. Training and testing results with different inputs

Input Vector	RMSE	VR	SCP
$X(t-k)$, $k = 0,1,\cdots 5$, only	0.57	50.80%	65%
With $Y_1(t)$ (S&P 500) added	0.54	56.85%	76%
With the Day of the Week added	0.53	57.26%	80%

According to the SCP scores, the US S&P 500 index and the day of the week have added 11% and 4 % additional information to the prediction based on the Australian stock market index alone.

6 Further Possibilities

So far we have only investigated very basic aspects of the problem. The results obtained are already very useful and the trained neural networks have been used routinely to generate daily predictions on the Australian stock index in our real-money stock and index futures trading. However, it should be pointed out that developing a specific neural network or a set of neural networks for this prediction purpose is a never-ending evolutionary process. Further possibilities for improving the correctness, accuracy and reliability of the prediction include selection of technical indicators, in particular, those based on wavelet and fractal analysis, prediction of marginal probability distribution, ensembles or committees of neural networks, profit-driven genetic training algorithms, finer representation of time signatures such as using complex numbers for the day of the week, the day or week of the month, the day or week or month of the year in the input vector, and so on.

7 Conclusions

A minimal neural network has been developed for predicting the Australian stock market index AORD, which takes the last 6 daily returns of the target market, the last daily return of US S&P 500 index, and the day of the week as the input vector. It has only one hidden layer of 2 hidden neurons. It has achieved correctness in directional prediction of 80%. In additional to this development, a 6-day cycle has been discovered in the Australian stock market for the first time. These results shed strong light on further research and development directions.

References

[1] Azoff E.M. (1994): Neural Network Time Series Forecasting of Financial Markets. Wiley.

[2] Baestaens D.E., van den Berg W.M. and Vaudrey H. (1995): Money market headline news flashes, effective news and the DEM/USD swap rate: An intraday analysis in operational time. *Proc. 3rd Forecasting Financial Markets Conference*, London.

[3] Cybenko G. (1988): Continuous valued neural networks with two hidden layers are sufficient. Technical report, Department of Computer Science, Tufts University, Medford, Massachusetts, USA.

[4] Cybenko (1989): Approximation by superpositions of a signoidal function. *Mathematics of Controls, Signals and Systems*, 2: 303-314.

[5] Gately (1996): Neural Networks for Financial Forecasting. Wiley.

[6] Hornik K., Sinchcombe M. and White H. (1989): Multilayer feedforward networks are universal approximators. *Neural Networks*, 2: 259-366.

[7] Murphy J. (1999): Technical Analysis of the Financial Markets: A Comprehensive Guide to Trading Methods and Applications. Prentice Hall Press.

[8] Pan H.P. (2003a): A joint review of technical and quantitative analysis of the financial markets towards a unified science of intelligent finance. Proc. *2003 Hawaii International Conference on Statistics and Related Fields*, June 5-9, Hawaii, USA.

[9] Pan H.P. (2003b): Swingtum – A computational theory of fractal dynamic swings and physical cycles of stock market in a quantum price-time space. Ibid.

[10] Pan H.P. and Foerstner W. (1992): An MDL-principled evolutionary mechanism to automatic architecturing of pattern recognition neural network. *IEEE Proc. 11th International Conference on Pattern Recognition*, the Hague.

[11] Qi M. (1999): Nonlinear predictability of stock returns using financial and economic variables. *Journal of Business and Economic Statistics*, 17: 419-429.

[12] Shih Y.L. (1991): Neural nets in technical analysis. *Technical Analysis of Stocks and Commodities*, 9(2):62-68.

[13] Fishman M.B., Barr D.S. and Loick W.J. (1991): Using neural nets in market analysis. *Technical Analysis of Stocks and Commodities*, 9(4): 18-22.

[14] Katz J.O. (1992): Developing neural network forecasters for trading. *Technical Analysis of Stocks and Commodities*, 10(4).

[15] Kean J. (1992): Using neural nets for intermarket analysis. *Technical Analysis of Stocks and Commodities*, 10(11).

[16] Refenes A.-P., Abu-Mostafa Y., Moody J. and Weigend A. (eds) (1996): *Neural Networks in Financial Engineering.* Proc. 3rd International Conference on Neural Networks in the Capital Markets. World Scientific.

[17] Rogers R. and Vemuri V. (1994): *Artificial Neural Networks Forecasting Time Series.* IEEE Computer Society Press, Los Alamitos, CA.

[18] Ruggerio M. (1994): Training neural nets for intermarket analysis. *Futures*, Sep: 56-58.

[19] Stein (1991): Neural networks: from the chalkboard to the trading room. *Futures*, May: 26-30.

[20] Swales G.S. and Yoon Y. (1992): Applying artificial neural networks to investment analysis. *Financial Analysts Journal*, 48(5).

[21] Utans J. and Moody J.E. (1991): Selecting neural network architectures via the prediction risk: application to corporate bond rating prediction. *Proc. 1st Int. Conference on AI Applications on Wall Street*, IEEE Computer Society Press.

[22] Virili F. and Reisleben B. (2000): Nonstationarity and data preprocessing for neural network predictions of an economic time series. *Proc. Int. Joint Conference on Neural Networks 2000*, Como, Vol. 5, pp. 129-136.

[23] Ward S. and Sherald M. (1995): The neural network financial wizards. *Technical Analysis of Stocks and Commodities*, Dec: 50-55.

[24] White H. (1988): Economic predictions using neural networks: the case of IBM daily stock returns. *Proc. IEEE International Conference on Neural Networks*, Vol. 2, pp. 451-458.

[25] Wong F.S. (1992): Fuzzy neural systems for stock selection. *Financial Analysts Journal*, 48: 47-52.

[26] Yao J.T. and Tan C.L. (2001): Guidelines for financial forecasting with neural networks. *Proc. International Conference on Neural Information Processing*, Shanghai, China, pp. 757-761.