EQUITY FORECASTING: A CASE STUDY ON THE KLSE INDEX

JINGTAO YAO HEAN-LEE POH
Department of Information Systems and Computer Science
National University of Singapore
Kent Ridge, Singapore 0511
Fax: 65 779 4580
E-mail: yaojingt@iscs.nus.sg pohhl@iscs.nus.sg

ABSTRACT

This paper presents the research of neural networks as applied in equity forecasting in an emerging market such as the Kuala Lumpur Stock Exchange(KLSE). Backpropagation neural networks are used to capture the relationship between the technical indicators and the levels of the KLSE index over time. The experiment shows that useful predictions can be made without the use of extensive market data or knowledge. In fact, a significant paper profit can be achieved by purchasing indexed stocks in the respective proportions. The paper, however, also discusses the problems associated with technical forecasting using neural networks, such as the choice of "time frames" and the "recency" problems.

KEY WORDS: Neural Network, Financial Analysis, Emerging Equity Market, Prediction

1. Introduction

Equity has long been considered a high return investment field. The major forecasting methods used in the financial area are either technical or fundamental. Due to the fact that stock markets are affected by many highly interrelated economic, political and even psychological factors, interaction among these indicators became very complex. Therefore, it is generally very difficult to forecast the movement in the stock market.

Classical statistical techniques for forecasting reach their limitation in applications with nonlinearities in the data set[11]. Neural networks have been shown to be able to decode nonlinear time series data which adequately describe the characteristics of stock markets[8].

Stock market prediction has been a research topic for many years[9]. There are two important hypotheses[9] for the possibility of forecasting: the *random walk hypothesis* and the *efficient market hypothesis*. The research done here would be considered a violation of these hypotheses for short term trading advantages in one of the most important emerging markets, which is considered by some Malaysian researchers such as Yong [13] [14] to be "less random" than the mature markets. In fact, the stock market price movements of United States[5] and Japan[1] have been shown to conform to the weak form of the efficient market hypothesis, which states that no investor can achieve trading advantages based on historical stock market price and volume information. Solnik[10] studied 234 stocks from eight major European stock markets and indicated that these European stock markets exhibited a slight departure from the random walk. However, all these studies were conducted at different points of time and in different time intervals. It is therefore hard to conclude that even the weak form of the efficient market hypothesis is supported in any market at all times and under all circumstances.

This paper begins with the general discussion on the possibilities of equity forecasting in an emerging market. It is followed by a section on neural networks. Subsequently, a section is devoted to a case study on the equity forecasting in one of the largest emerging markets, pointing to the promises and problems of such an experiment. Finally, a conclusion which also discusses areas for future research is included at the end of the paper.

2. Neural Network and Equity Forecasting

Nowadays, traders no longer rely on a single indicator to provide information about the future trends of markets. They use a variety of indicators to obtain multiple signals. Neural networks are often trained by both technical and fundamental indicators to produce trading signals. Fundamental and technical analysis could be simulated in neural networks. For fundamental methods, retail sales, gold price, industrial production index, and foreign currency exchange rate etc. could be used as inputs[7]. For technical methods, the delayed time series data together with the technical indicators such as moving averages, stochastics, relative strength index etc. could be used as inputs. In this paper, backpropagation neural networks[12] are used to predict the daily prices of Kuala Lumpur Stock Exchange Composite Index (KLSE in short).

A neural network is a collection of interconnected simple processing elements. Every connection of neural network has a weight attached with it. The units in the network are connected in a feedforward manner, from the input layer to the output layer. The weights of connections have been given initial values. The error between the predicted output value and the actual value is backpropagated through the network for the updating of the weights. This is a supervised learning procedure that attempts to minimize the error between the desired and the predicted outputs. The output value for a unit j is given by the following function:

$$O_j = G(\sum_{i=1}^{m} w_{ij}x_i - \theta_j) \tag{1}$$

Where x_i is the output value of the ith unit in a previous layer, w_{ij} is the weight on the connection from the ith unit, θ_j is the threshold, and m is the number of units in the previous layer. The $G()$ is a sigmoid function:

$$G(z) = \frac{1}{1+e^z} \tag{2}$$

The $G()$ is a commonly used activation function for time series prediction in backpropagation networks[4 8].

A typical backpropagation neural network is used to capture the relationship between the stock prices of today and the future. The relationship can be obtained through a group of mappings of constant time interval. Assume that u_i represents today's price, v_i represents the price after ten days. If the prediction of a stock price after ten days could be obtained using today's stock price, then there should be a functional mapping from u_i to v_i, where

$$v_i = \Gamma_i(u_i) \tag{3}$$

Using all $(u_i,\ v_i)$ pairs of historical data, a general function $\Gamma()$ which consists of $\Gamma_i()$ could be obtained.

$$v = \Gamma(u) \tag{4}$$

More generally, $\vec{u}$ which consists of more information in today's price could be used in function $\Gamma()$. Neural networks can simulate all kinds of functions, so they also can be used to simulate this $\Gamma()$ function. The $\vec{u}$ is used as the inputs of the neural network.

There are five major steps in the neural network based forecasting. First, the information that could be used as the inputs and outputs of neural network is collected. Second, these data are normalized and scaled in order to reduce the fluctuation and noise. Third, a neural network model that could be used to capture the relationship between the data of inputs and outputs is built. Fourth, variations of the models, i.e., different models and configurations with different training, validation and testing data sets are invested. Finally, the best model measured by out-of-sample hit rates, for example, is chosen for use in forecasting.

3. A Case Study on the Forecasting of the KLSE Index

The KLSE is calculated on the basis of 86 major Malaysian stocks. It is capitalization-weighted by paäsche formula and has a base level of 100 as of 1977. It may be regarded as the Malaysian Dow Jones Index. As of March 13, 1995, 492 companies have been listed. It has only ten years of history, so there are not enough fundamental data that could be used for forecasting. Besides, the KLSE is considered a young and speculative market, where investors tend to look at price movements, rather than the fundamentals.

Hence, in this paper, a mixed technical method was adopted which takes not only the delayed time series data as inputs but also the technical indicators. Neural networks are trained to approximate the thinking and behavior of some stock market traders. Different indicators are used as the inputs to a neural network and the index of stock is used to supervise the training process, in order to discover implicit rules governing the price movement of KLSE. Finally, the trained neural network is used to predict the future levels of the KLSE index.

The technical analysis method is used commonly to forecast the KLSE index, the buying and selling point, turning point, and the highest, lowest point etc. When forecasting by hand, different charts will be used by analysts in order to predict the changes of stocks in the future. Neural networks could be used to recognize the patterns of the chart and the value of index.

3.1. Data Choice and Pre-processing

The daily data from Jan 3, 1984 to Oct 16, 1991(1911 data sets) are used on the first trial. These data are preprocessed to be used in the prediction of the daily indices.

Technical analysts usually use indicators to predict the future. The major types of indicators are *moving average(MA), momentum(M), Relative Strength Index(RSI) and stochastics(%K), and moving average of stochastics(%D).* These indicators can be derived from the real stock composite index. The target for training the neural network is the actual index.

The inputs of the neural network model are $I_{t-1}, I_t, MA_5, MA_{10}, MA_{50}$, RSI, M, %K and %D. The output is I_{t+1}. Here I_t is the index of *t-th* period, MA_j is the moving average after *j-th* period, and I_{t-1} is the delayed time series. For daily data, the indicators are calculated as mentioned above. Other indicators are defined as follows,

$$M = CCP - OCP \tag{5}$$

where
CCP = Current closing price
OCP = Old closing price for a predetermined period (5 days)

$$RSI = 100 - \frac{100}{1 + \frac{\Sigma(Positive\ Changes)}{\Sigma(Negative\ Changes)}} \tag{6}$$

$$\%K = \frac{CCP - L9}{H9 - L9}100 \tag{7}$$

where
L9 = the lowest low of the past 9 days

H9 = the highest high of the past 9 days

$$\%D = \frac{H3}{L3} 100 \tag{8}$$

where
H3 = the three day sum of (CCP - L9)
L3 = the three day sum of (H9 - L9)

Indicators can help traders identify trends and turning points. Moving average is a popular and simple indicator for trends. Stochastic and RSI are some simple indicators which help traders identify turning points.

In general, the stock price data have bias due to differences in name and spans. Normalization can be used to reduce the range of the data set to values appropriate for inputs to the activation function being used. The normalization and scaling formula is:

$$y = \frac{2x - (max - min)}{max + min} \tag{9}$$

Where
x is the data before normalizing
y is the data after normalizing.

Because the index prices and moving averages are in the same scale, so the same maximum and minimum data are used to normalize them. The *max* is derived from the maximum value of the linked time series, and the same applies to the minimum.

3.2. Measurements of Neural Network Training

The *Normalized Mean Squared Error (NMSE)* is used as one of the measures to decide which model is better. It can evaluate and compare the predictive power of the models. The definition of *NMSE* is:

$$NMSE = \frac{\sum_k (x_k - \hat{x}_k)^2}{\sum_k (x_k - \bar{x}_k)^2} \tag{10}$$

where the x_k and $\hat{x}_k$ represent the actual and predicted values respectively, and the $\bar{x}_k$ is the mean of x_k. Other evaluation measures include the calculation of the correctness of signs and gradients. Sign statistics can be expressed as

$$Sign = \frac{\sum s_k}{N} \tag{11}$$

where N represents the number of patterns in a testing set and s_k is a segment function which can be expressed as:

$$s_k = \begin{cases} 1 & x_k \hat{x}_k > 0 \\ 1 & x_k = \hat{x}_k = 0 \\ 0 & \text{otherwise} \end{cases} \tag{12}$$

Here $Sign$ represents the correctness of signs after normalization. Similarly, directional change statistics can be expressed as

$$Grad = \frac{\sum g_k}{N} \tag{13}$$

where

$$g_k = \begin{cases} 1 & (\hat{x}_{k+1} - x_k) = 0 \\ 1 & (x_{k+1} - x_k)(\hat{x}_{k+1} - x_k) > 0 \\ 0 & \text{otherwise} \end{cases} \tag{14}$$

$NMSE$ is one of the most popularly used measurements. It represents the fit between neural network predictions and actual targets. However, a prediction that follows closely the trend of the actual target would also result in a low $NMSE$. For pattern recognition, it is a very important signal. But for trading in the context of time series analysis, this may not be the case. The study done by the authors show that $Grad$ is a better indicator of the quality of forecasting for trading purposes. $Grad$ can show how good the forecast trend is, which is very useful for trading purposes. Similarly, when the inputs are the changes of levels instead of the actual levels, then $Sign$ also points to the accuracy of the forecast trends, and therefore it could be useful for trading.

3.3. Nonlinear Analysis of the KLSE Data

The mean, variance and standard deviation of KLSE index are 383.73, 15045.14 and 122.66 respectively. Figure 1 shows the graph of the KLSE index represented as logarithmic return $ln(I_{t+1}/I_t)$ for the defined period, where I_t is the index value at time t. It shows that the data is very noisy which makes forecasting very difficult.

The rescaled range analysis(R/S analysis)[6] is able to distinguish a random series from a non-random series, irrespective of the distribution of the underlying series. In this paper, it is used to detect the long-memory effect in the KLSE time series over a time period. R can capture the maximum and minimum

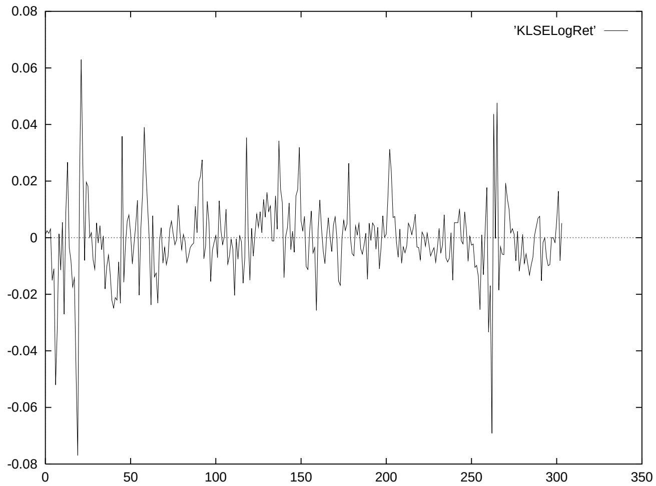

Figure 1: Logarithmic Returns of KLSE Daily Data From July 30, 1990 (horizontal scale 0) to Oct 16, 1991(304)

cumulative deviations of the observations x_t of the time series from its mean (μ), and it is a function of time (the number of observations N):

$$R_N = \max[x_{t,N}] - \min[x_{t,N}] \tag{15}$$

where

$$x_{t,N} = \sum_{t=1}^{N}(x_t - \mu), \qquad t = 1, .. \tag{16}$$

The R/S ratio of R and the standard deviation S of the original time series can be estimated by the following empirical law: $R/S = N^H$ when observed for various N values. For some values of N, the Hurst exponent can be calculated by:

$$H = log(R/S)/log(N), \qquad 0 < H < 1. \tag{17}$$

and the estimate of H can be found by calculating the slope of the *log/log* graph of R/S against N using regression.

The value of Hurst exponent was found be 0.881444 which denotes long-memory effect in time series. Hence, there exist possibilities for conducting time series forecasting in the KLSE data.

3.4. Model Building

Usually one third of the collected data is used for testing and two thirds of the data for training. Two fifths of the testing data is used for validation. A model is considered good if the error for out-of-sample testing is the lowest compared with the other models. If the trained model is the best one for validation and also the best one for testing, one can assume that it is a good model for future forecasting.

The data are chosen and segregated in time order. In other words, the data of the earlier period are used for training, the data of the later period are used for validation, and the data of the latest time period are used for testing. This method may have some *recency problems*. Using the above rule, the neural network is only trained using data up till the end of 1988. In forecasting the index after November 1991, the neural network is 'forced' to use knowledge up till 1988 only. Hence, another method where the data are randomly chosen is designed to circumvent this problem.

After experimenting with the choice of data, a very good testing result may not predict well. On the other hand, a model which is trained with randomly chosen data may predict well even with average testing results.

The main consideration when building a suitable neural network for the financial application is to make a trade-off between convergence and generalization. It is important not to have too many nodes in the hidden layer because this may allow the neural network to learn by example only and not to generalize[2].

According to Beale and Jackson[3], a network with one hidden layer can model any continuous function. Depending on how good we want to approximate our function, we may need tens, hundreds, thousands, or even more neurons. In practice, people would not use one hidden layer with a thousand neurons, but prefer more hidden layers with fewer neurons doing the same job. Of course there is probably no perfect rule of thumb and the best thing to do is to use a cross validation set to obtain the optimal generalization point. An iterative process is adopted beginning with one node in a hidden layer and working up until a minimum error in the test set is obtained.

The models constructed have configurations such as 5-3-1, 5-4-1, 5-6-1, 5-3-2-1, 6-3-1, 6-5-1, 6-5-1, and 6-4-3-1. For the five inputs, I_t, MA_5, MA_{10}, RSI and M are used. For the six inputs, $I_{t-1}, I_t, MA_5, MA_{10}, RSI$ and M are used. The best results for different models using daily data are shown in the Table 1. The %K and %D indicators have been found to be less sensitive for training in this case. The five-input-models and six-input-models show the best results in the experiments.

Figure 2 is the result of the prediction using out-of-sample testing data.

Using NMSE as a measure, the 5-3-2-1 model is the best. However, the 5-4-1 model delivers the best gradient prediction correctness, and the 6-4-3-1 model

Arch	α	η	Max Err	NMSE	Grad	Sign
5-3-1	0.005	0.0	0.001	0.231175	67%	77%
5-4-1	0.005	0.0	0.001	0.178895	**85%**	86%
5-3-2-1	0.005	0.1	0.001	**0.032277**	78%	83%
6-3-1	0.005	0.1	0.001	0.131578	82%	66%
6-5-1	0.005	0.0	0.001	0.206726	78%	89%
6-4-3-1	0.005	0.1	0.001	0.047866	75%	**96%**

Table 1: The Best Results for Six Different Models (*Arch*: architecture of the neural network; α: learning rate; η: momentum rate; *Max Err* : maximum error for one pass thru the whole training set; NMSE: normalized mean squared error; *Grad*: correctness of gradients; and *Sign*: correctness of signs.)

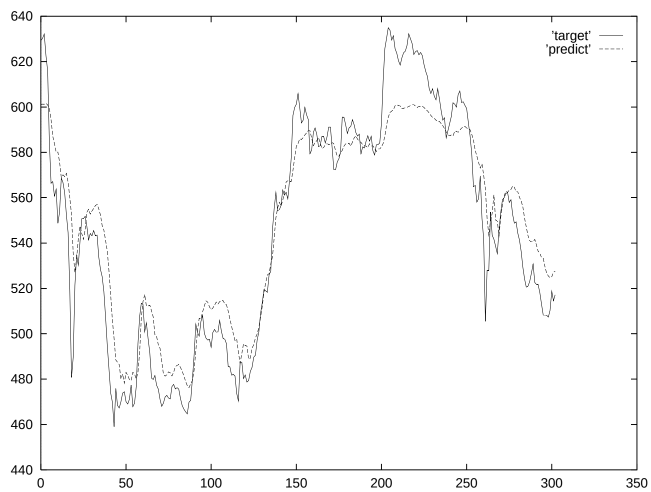

Figure 2: Daily Stock Price Index Prediction of KLSE (Out of Sample Data: From July 30, 1990(horizontal scale 0) to Oct 16, 1991(304))

gives the best prediction in terms of sign.

3.5. Limitations and Trade Offs

A very small NMSE does not necessarily imply good generalization. In fact, the choice of the testing criteria, be it NMSE, or sign, or gradient, depends on the trading strategies. In the stock market, the gradient is very important for traders. Sometimes, they even need not know what the actual level of the index is.

The sum of the NMSE of the three parts of data (training, validation and testing) must be kept small, not just the training NMSE. Sometimes having small NMSEs for testing and validation is more important than having small

Architecture	Return-1	Return-2
5-3-1	38.422423%	9.036855%
5-4-1	40.137030%	11.906852%
5-3-2-1	48.887812%	22.943906%
6-3-1	42.481686%	12.740843%
6-5-1	36.475363%	10.237681%
6-4-3-1	47.046146%	**26.023073%**

Table 2: Paper Profit for Different Models (Return-1: Annual Return using Entire Data Set, Return-2: Annual Return using Out of Sample *Daily* Data Set)

NMSE for training.

Further, better testing results are demonstrated in the period near the end of the training sets. This is a result of the 'recency' problem. There should be some recency trade offs. One shouldn't say it's the best model unless one has tested it, but once one has tested it one has not trained the model enough. Also, the hit rate depends on the time frame chosen for out-of-sample testing. Generalization of the predictability of the market trends to other time period is still difficult. In other words, the neural network model may work well in forecasting the market trends in certain period of time only.

3.6. Simulated Profits

As mentioned earlier, the evaluation of the model depends on the strategy of the traders or investors. To simulate these strategies, a small program was developed. Assume that a certain amount of seed money is used in this program. The seed money is used to buy a certain amount of indexed stocks in the right proportion when the prediction shows a rise in the stock price. To calculate the profit, the major blue chips in the KLSE index basket are bought or sold at the same time. To simplify the calculation, assume that the aggregate price of the major blue chips is the same as the KLSE index. A method used is to buy the stock with the price of index from the point that neural network predicts the stock price will rise. And then the 'stock' will be held at hand till the next turning point that neural network predicts. The results obtained are shown in Table 2. The best return based on the out of sample prediction is obtained with the 6-4-3-1 model using the above simulation method. The annual return rate is approximately 26% . If the prediction generated by the entire training and testing data is used, then the annual return rate would be approximately 47% .

Transaction cost will be considered in real trading. In this paper, 1% of transaction cost was included in the calculation. In some stock markets, the indices are traded in the derivative markets. However, the KLSE index is not traded anywhere. Therefore the indexed stocks were bought or sold in proportional amounts in this paper. In a real situation, this might not be possible as some indexed stocks may not be traded at all on some days. Besides, the transaction cost of a big fund trading, and thus affecting the market prices was not taken into consideration in the calculation of the "paper profit" or simulated profit. To be more realistic, a certain amount of transaction cost would have to be included in the calculation.

There are two benchmarks for the simulated profit. *Benchmark 1* is passive investment method which uses the seed money to buy the index on the first day of testing period(July 30, 1990) and sell it on the last day of this period(Oct 16, 1991). The annual return for *Benchmark 1* is −14.98%. *Benchmark 2* is to save the seed money at the beginning and withdraw it at the end earning interest. The annual return for *benchmark 2* is 7.98%. The two returns are calculated as follows:

$$Return\ 1 = \left(\frac{index2}{index1}\right)^{\frac{12}{n}} - 1 \tag{18}$$

where
index1 = index on first testing day
index2 = index on last testing day
n= No. of months in testing period

$$Return\ 2 = \left[\prod_{j=Jul\ 90}^{Oct\ 91}\left(1 + \frac{int_j}{12}\right)\right]^{\frac{12}{n}} - 1 \tag{19}$$

where
int_j = interest rate in jth month of testing period

4. Conclusion

The performance of several backpropagation neural networks applied to the problem of predicting the KLSE stock market index was evaluated. The delayed index levels and some technical indicators were used as the inputs of neural networks, while the current index level was used as output. With the prediction, significant paper profits were obtained for a chosen testing period of 304 trading days in 1990/91. Besides, the experiments showed that useful prediction could be made without the use of extensive market data or knowledge.

There are four challenges beyond the choice of either technical or fundamental data for using neural network to forecast the stock prices. First, the inputs and outputs of the neural networks have to be determined and preprocessed. Second, the types of neural networks and the activation functions for each node have to be chosen. Third, the neural network architecture based on the experiment with different models has to be determined. Finally, different measures to evaluate the quality of trained neural networks for forecasting have to be experimented with.

The same mixed technical analysis was also applied to weekly data. However, the results were not as impressive as those obtained using the daily data. This is attributed to the high volatility of the KLSE market.

The significance of this research is as follows:

- It showed how a 26% annual return could be achieved by using the discussed model. The annual return for passive investment and bank saving were -14.98% and 7.98% respectively for same period of time. Thus, the results showed that there is practical implication for an index linked fund to be set up in the cash market or for the index to be traded in the derivative market.

- It highlighted the problems associated with neural network based time series forecasting. Such problems are, for instance, the fact that

  - the hit rate is a function of the time frame chosen for the testing sets;
  - generalizability of the model over time to other period is weak;
  - there should be some recency trade offs.

For future research, a comparison with conventional forecasting methods should be made to assess the efficacy of neural networks in time series forecasting. To improve neural networks' capabilities in forecasting, a mixture of technical and fundamental factors as inputs over different time periods should be considered. Sensitivity analysis should be conducted which can provide pointers to the refinement of neural network models. Last but not least, the characteristics of emerging markets such as KLSE should be further researched to facilitate better modeling of the market using neural networks. The forecasting results can then be applied to the trading of index linked stocks under consideration of transaction costs.

5. References

1. Ang, J.S. and Pohlman, R. A., "A Note on the Price Behavior of Far Eastern Stock." *Journal of International Business Studies*, Spring/Summer, 1978
2. Baum, E. B. and Hassler, D. , "What Size Net Gives Valid Generalization? " *Neural Computation*, **1**, (1989) 151-160.
3. Beale, R. and Jackson, T., *Neural Computing: An Introduction*, Adam Hilger, 1990
4. Chapman, A. J., "Stock Market Reading Systems Through Neural Networks: Developing a Model", *Int. Journal of Applied Expert Systems* **Vol. 2, no. 2,** 1994 88-100
5. Fama, E. F., "The Behavior of Stock Market Prices", *Journal of Business* Jan. 1965 34-105
6. Hurst, H. E., "Long Term Storage of Reservoirs", *Transactions of the American Society of Civil Engineers 116,* 1951
7. Jasic, T. and Poh, H. L., "Financial Time-Series Prediction Using Neural Networks: A Case Study for the TOPIX Data", *Proceedings of the Sixth Australian Conference on Neural Networks-ACNN'95,* 6-8 February 1995, Sydney, Australia
8. Lapedes, A. and Farber, R., "Nonlinear Signal Processing Using Neural Networks," *IEEE Conference on Neural Information Processing System - Natural and Synthetic*, 1987
9. Peters, E. E., *Chaos and Order in the Capital markets: A New View of Cycles, Prices, and Market Volatility*, John Wiley & Sons Inc. 1991
10. Solnik, B. H., "Note on the Validity of the Random Walk for European Stock Prices." *Journal of Finance* Dec. 1973
11. Refenes, A. N., Zapranis, A. and Francis, G., "Stock Performance Modeling Using Neural Networks: A Comparative Study with Regression Models", *Neural Network*, **Vol 5** 1994, 961-970
12. Rumelhart, D.E. et al, "Learning Internal Representations by Error Propagation", *Parallel Distributed Processing: Explorations in the macrostructure of cognition,* **Vol. 1**, MIT Press
13. Yong, O., "Behavior of the Malaysian Stock Market" Penerbit Universiti Kebangsaan Malaysia, 1993, Malaysia
14. Yong, O., "The Malaysian Stock Market and You" Leeds Publication, 1995, Malaysia

Paper of NNCM'95(3rd International Conference On Neural Networks in the Capital Markets), Oct 1995, The Langham Hilton, London, England.