The Korean Communications
in Statistics Vol. 8, No. 2, 2001
pp. 543-556

Neural Network Forecasting Using Data Mining Classifiers Based on Structural Change: Application to Stock Price Index

Kyong Joo Oh[1] and Ingoo Han[2]

Abstract

This study suggests integrated neural network models for the stock price index forecasting using change-point detection. The basic concept of this proposed model is to obtain significant intervals occurred by change points, identify them as change-point groups, and reflect them in stock price index forecasting. The model is composed of three phases. The first phase is to detect successive structural changes in stock price index dataset. The second phase is to forecast change-point group with various data mining classifiers. The final phase is to forecast the stock price index with backpropagation neural networks. The proposed model is applied to the stock price index forecasting. This study then examines the predictability of integrated neural network models and compares the performance of data mining classifiers.

Keywords : Structural Change, Change-Point Detection, Pettitt Test, Backpropagation Neural Networks

1. Introduction

Stock markets are incredible systems wherein thousands of instruments are traded by millions of participants every day, around the world, in a never-ending battle to make money. Previous studies on stock market prediction using artificial neural networks (ANN) have been executed during the past decades. These studies used various types of ANN to predict the stock price index and the direction of its change. Previous work in stock market prediction has tended to use statistical techniques and artificial intelligence (AI) techniques in isolation. However, an integrated approach, which makes full use of statistical approaches and AI techniques, offers the promise of better performance than each method alone.

---

1) Researcher, Techno-Management Institute, Korea Advanced Institute of Science and Technology, Seoul 130-012, Korea.
E-mail : johanoh@kgsm.kaist.ac.kr
2) Associate Professor, Graduate School of Management, Korea Advanced Institute of Science and Technology, Seoul 130-012, Korea.
E-mail : ighan@kgsm.kaist.ac.kr

The early days of these studies focused on estimating the level of return on stock price index. Kimoto et al. (1990), one of the earliest studies for stock market prediction using AI, employed several learning algorithms and prediction methods for the Tokyo stock exchange prices index (TOPIX) prediction system. Their system used modular neural networks to learn the relationships among various factors. Kamijo and Tanigawa (1990) used recurrent neural networks for analyzing candlestick charts. Ahmadi (1990) used backpropagation neural networks with the generalized delta rule to predict the stock market. They intended to test the *Arbitrage Pricing Theory* (APT) using ANN. Yoon and Swales (1991) also performed predictions using qualitative and quantitative data. Some researchers investigated the issue of predicting the stock index futures market. Trippi and DeSieno (1992) and Choi et al. (1995) predicted the daily direction of change in the S&P 500 index futures using ANN. Duke and Long (1993) executed the daily predictions of the German government bond futures using feedforward backpropagation neural networks.

Recent research tends to include novel factors and to hybridize several AI techniques. Hiemstra (1995) proposed fuzzy expert systems to predict stock market returns. He suggested that ANN and fuzzy logic could capture the complexities of functional mapping because they do not require the specification of the function to approximate. A more recent study of Kohara et al. (1997) incorporated prior knowledge to improve the performance of stock market prediction. Tsaih et al. (1998) integrated the rule-based technique and ANN to predict the direction of the S&P 500 stock index futures on a daily basis. In this study, we suggest the integrated neural network model based on the statistical change-point detection.

In general, macroeconomic time series data is known to have a series of change points since they are controlled by governments monetary policy (Mishkin, 1995; Oh and Han, 2000). However, previous studies did not consider the structural break or the change-point in stock price index forecasting. The government takes intentional action to control the currency flow that has direct influence upon fundamental economic indices. For the stock price index, institutional investors play a very important role in determining its ups and downs since they are major investors in terms of marking and volume for trading stocks. They respond sensitively to such economic indices like stock price indices, the consumer price index, anticipated inflation, etc. Therefore, we can conjecture that the movement of the stock price index also has a series of change points.

Based on these inherent characteristics in stock price index, this study suggests the integrated neural network model using change-point detection, which plays the role of clustering the time series dataset. The proposed model consists of three phases: the first phase is to detect successive change points in the stock price index dataset, the second phase is to forecast the change-point group with backpropagation neural networks (BPN), and the final stage is to forecast the output with BPN. This study then examines the predictability of the proposed model. To explore the predictability, we divided the stock price index data into the training data over one period and the testing data over the next period. The predictability of the proposed model is examined using the metrics of the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE).

We outline the development of change-point detection and its application to the financial economics in Section 2. Section 3 describes the proposed integrated neural network model details through the various data mining classifiers. Section 4 reports the processes and the results of applied study. Finally, the concluding remarks are presented in Section 5.

2. Change-Point Detection

2.1. Existence and Detection of Structural Change for the Financial Economics

The detection and estimation of a structural or parametric change in forecasting is an important and difficult problem. In particular, financial analysts and econometricians have frequently used piecewise-linear models which also include change-point models. They are known as models with structural breaks in the economics literature. In these models, the parameters are assumed to shift - typically once - during a fixed sample period and the goal is to estimate the two sets of parameters as well as the change point or structural break.

In order to detect the structural change, change-point detection methods have been applied to macroeconomic time series. Rappoport and Reichlin (1989) and Perron (1989, 1990) conduct the first study in this field. From then on, several statistics have been developed which work well in a change-point framework, all of which are considered in the context of breaking the trend variables (Banerjee et al., 1992; Christiano, 1992; Zivot and Andrews, 1992; Perron, 1995; Vogelsang and Perron, 1995). In those cases where only a shift in the mean is present, the statistics proposed in the papers of Perron (1990) or Perron and Vogelsang (1992) stand out.

In spite of the significant advances by these works, we should bear in mind that some variables do not show just one change point. Rather, it is common for them to exhibit the presence of multiple change points. Thus, it seems advisable to introduce a large number of change points in the specifications of the models that allow us to obtain the abovementioned statistics. For example, Lumsdaine and Papell (1997) have considered the presence of two or more change points in trend variables. Based on this fact, we also assume the stock price index have two or more change points in our research model.

Up to date, there are few artificial intelligence models for financial applications to represent the change-point detection problems. Most of the previous research has a focus on the finding of unknown change points for the past, not to forecast for the future (Wolkenhauer and Edmunds, 1997; Li and Yu, 1999). Our model finds change points in the learning phase and forecasts change points in the testing phase. It is demonstrated that the introduction of change points to our model will make the predictability of stock price index greatly improve.

2.2. the Pettitt Test

In this study, a series of change points will be detected by the Pettitt test (Pettitt, 1979, 1980a), a nonparametric change-point detection method, since nonparametric statistical property is a suitable match for a neural network model that is a kind of nonparametric method (White, 1992). In this point, the introduction of the Pettitt test is fairly appropriate for the analysis of chaotic time series data. The Pettitt test is explained as follows.

Consider a sequence of random variables $X_1, X_2, \dots, X_T$, then the sequence is said to have a change-point at τ if X_t for $t=1, 2, \dots, \tau$ have a common distribution $F_1(x)$ and X_t for $t=\tau+1, \tau+2, \dots, T$ have a common distribution $F_2(x)$, and $F_1(x) \neq F_2(x)$. We consider the problem of testing the null hypothesis of *no-change*, $H_0: \tau = T$, against the alternative hypothesis of *change*, $H_A: 1 \leq \tau < T$, using a nonparametric statistic.

An appealing non-parametric test to detect a change would be to use a version of the Mann-Whitney two-sample test. A Mann-Whitney type statistic has remarkably stable distribution and provides a robust test of the change point resistant to outliers (Pettitt, 1980b). Let

$$D_{ij} = sgn(X_i - X_j) \tag{1}$$

where $sgn(x)=1$ if $x>0$, 0 if $x=0$, -1 if $x<0$, then consider

$$U_{t,T} = \sum_{i=1}^{t} \sum_{j=t+1}^{T} D_{ij}. \tag{2}$$

The statistic $U_{t,T}$ is equivalent to a Mann-Whitney statistic for testing that the two samples $X_1, X_2, \dots, X_t$ and $X_{t+1}, X_{t+2}, \dots, X_T$ come from the same population. The statistic $U_{t,T}$ is then considered for values of t with $1 \leq \tau < T$. For the test of H_0: *no change* against H_A: *change*, we propose the use of the statistic

$$K_T = \max_{1 \leq t < T} |U_{t,T}|. \tag{3}$$

The limiting distribution of K_T is $\Pr \cong 2\exp\{-6k^2/(T^2+T^3)\}$ for $T \to \infty$.

In the time sequence dataset, the Pettitt test detects a possible change point in which the structural change is occurred. Once the structural change is detected through the test, the dataset is divided into two intervals. The intervals before and after the change point form homogeneous groups which take heterogeneous characteristics from each other. This process becomes a fundamental part of the binary segmentation method explained in Section 3.

3. Model Specification

3.1. Integrated Neural Network Model based on the Structural Change

Data mining classifiers, change-point detection method and neural network learning methods have been integrated to forecast stock price index. The advantages of combining multiple techniques to yield synergism for discovery and prediction have been widely recognized (Gottman, 1981; Kaufman et al., 1991). The proposed models are determined by the kind of data mining classifier which is applied to the second phase of model. This section provides the architecture and the characteristics of our model to involve the change-point detection and the BPN. Based on the Pettitt test, the proposed model is composed of three phases as follows:

Phase 1: Constructing homogeneous groups

Pettitt test is a method to find a change-point in longitudinal data (Pettitt, 1979). It is known that stock price index at time t are more important than fundamental economic variables in determining stock price index at time $t+1$ (Larrain, 1991). Thus, we apply Pettitt test to stock price index at time t to generate a forecast for $t+1$ in the learning phase. We, first of all, have to decide the number of change point. If change point is assumed to occur just one in given dataset, only the first step will be performed. Otherwise, all of three steps will be performed successively. The interval made by this process is defined as the significant interval, labeled SI, which is idenfied with a homogeneous group.

Step 1: Find a change point in $1 \sim N$ intervals by the Pettitt test. If r_1 is a change point, $1 \sim r_1$ intervals are regarded as SI_1 and $(r_1+1) \sim N$ intervals are regarded as SI_2. Otherwise, it is concluded that a change point does not occur for $1 \sim N$ intervals. $(1 \leq r_1 \leq N)$

Step 2: Find a change point in $1 \sim r_1$ intervals by the Pettitt test. If r_2 is a change point, $1 \sim r_2$ intervals are regarded as SI_{11} and $(r_2+1) \sim r_1$ intervals are regarded as SI_{12}. Otherwise, $1 \sim r_1$ intervals are regarded as SI_1 like Step 1. $(1 \leq r_2 \leq r_1)$

Find a change point in $(r_1+1) \sim N$ intervals by the Pettitt test. If r_3 is a change point, $(r_1+1) \sim r_3$ intervals are regarded as SI_{21} and $(r_3+1) \sim N$ intervals are regarded as SI_{22}. Otherwise, $(r_1+1) \sim N$ intervals are regarded as SI_2 like Step 1. $(r_1 \leq r_3 \leq N)$

Step 3: By applying the same procedure of Step 1 and 2 to subsamples, we can obtain several significant intervals under the dichotomy if we need five or more significant intervals.

First of all, the number of structural change should be determined. If just one change point is assumed to occur in a given dataset, only the first step will be performed. Otherwise, all of the three steps will be performed successively. This process plays a role of clustering that constructs groups as well as maintains the time sequence. In this point, Phase 1 is distinguished from other clustering methods such as the k-means nearest neighbor method and the hierarchical clustering method. They classify data samples by the Euclidean distance between cases without considering time sequence.

Phase 2: Group forecasting with data mining classifier

The significant intervals by Phase 1 are grouped to detect the regularities hidden in them and to represent the homogeneous characteristics of them. Such groups represent a set of meaningful trends encompassing the significant intervals. Since those trends help to find regularity among the related output values more clearly, the neural network model can have a better ability of generalization for the unknown data. This is indeed a very useful point for sample design. In general, the error for forecasting may be reduced by making the subsampling units within groups homogeneous and the variation between groups heterogeneous (Cochran, 1977). After Phase 1 detects the appropriate groups hidden in the significant intervals, various classifiers are applied to the input data samples at time t with group outputs for $t+1$. In this sense, Phase 2 is a model that is trained to find an appropriate group for each given sample.

Phase 3: Forecasting the output with BPN

Phase 3 is built by applying the BPN model to each group. Phase 3 is a mapping function between the input sample and corresponding desired output (i.e. stock price index). Once Phase 3 is built, then the sample can be used to forecast the stock price index.

3.2. The proposed models

According to the kind of classifier used in the Phase 2, we propose three integrated neural network models: (1) MDA (multivariate discriminant analysis)-supported neural network model, (2) CBR (case based reasoning)-supported neural network model and (3) BPN (backpropagation neural networks)-supported neural network model.

3.2.1. MDA-supported neural network model

MDA is used to classify individuals into one of two or more alternative groups on the basis of a set of measurements. Theoretically, This method is based on the Fisher's linear discriminant function by maximizing the ratio of between-groups and within-groups variances (Fisher, 1936). The groups are known to be distinct, and each individual belongs to one of them. The MDA can be used to identify which variables contribute to making the classification (McLachlan, 1992; Hair, et al., 1995). This model applies MDA to forecast the change-point group in the second stage.

3.2.2. CBR-supported neural network model

CBR can be an effective tool for predicting temporal variables. By extending the concept of a case to include several immediate neighbors and by seeking multiple exemplars from the case base, a CBR system can forecast a trajectory with a remarkable accuracy. In fact, such a system can exceed the performance of neural networks in some events, a methodology which has been widely applied to the task of forecasting (Deboeck and Cadar, 1994; Trippi and Turban, 1992; Kolodner, 1991, 1993). This model uses CBR to forecast the change-point group in the second stage.

3.2.3. BPN-supported neural network model

The neural network methodology has been applied extensively to solve practical problems following the publication of the backpropagation algorithm for the multi-layer perceptron (Rumelhart, 1986). The algorithm was developed for the perceptron model, a simple structure to simulate a neuron (Rosenblatt, 1957). Today the BPN is the most widely used neural networks algorithm in science, engineering, finance and other fields. This model uses the BPN two times. In the second stage, BPN is applied to forecast the change-point group. Finally, BPN also forecast the desired output.

4. Empirical Results

Research data used in this study comes from the weekly Korea Stock Price Index (KOSPI) from January 1987 to August 1996. The total number of samples includes 502 trading weeks. Input variables are 13 technical indicators. These variables are selected based on the review of domain experts. The description of input variables is presented in Table 1 (Achelis, 1995; Chang et al., 1996; Choi, 1995; Edwards and Magee, 1997; Gifford, 1995).

The training phase involves observations from January 1987 to December 1992 while the testing phase runs from January 1993 to August 1996. The stock price index data is presented in Figure 1. Figure 1 shows that the movement of stock price index highly fluctuates.

The Pettitt test is applied to the stock price index data. In this study, KOSPI data is varied from just one change-point to three change-points. For more change points, the proposed model can be applied by Phase 1. The study employed the regression model with autoregressive error term. labeled Pure_TS, and two kinds of neural network models. The first type of neural network models, labeled Pure_NN, involves thirteen input variables (shown at Table 1) at time t to generate a forecast for $t+1$. The second type has two-step forecasting models which consist of three phases mentioned in Section 3. The first step is Phase 2 that forecasts the change-point group while the next step is Phase 3 that forecasts the output. The classifiers used in the model were as shown in Table 2. For validation, five learning models were also compared.

Table 1. Description of input variables (C_t: Closing price at time t)

Variable Name	Formular
MA6 (6-day moving average)	$\frac{1}{6}\sum_{i=1}^{6} C_{t-i+1}$
RSI (Relative strength index)	$100 - \dfrac{100}{1 + \dfrac{\sum_{i=1}^{6} UP_{t-i+1}/6}{\sum_{i=1}^{6} DOWN_{t-i+1}/6}}$ where UP: Upward price change, and $DOWN$: Downward price change
OBV (On balance volume)	$OBV_{t-1} + kV_t$ $k=1$ if $C_t > C_{t-1}$, -1 if $C_t < C_{t-1}$, 0 if $C_t = C_{t-1}$ where V_t means the volume at time t
%K (Stochastic %K)	$\dfrac{C_t - LL_{t-6}}{HH_{t-6} - LL_{t-6}} \times 100$ where HH_n and LL_n mean the highest high and the lowest low in the last n days respectively
%D (Stochastic %D)	$\sum_{i=1}^{6} \%K_{t-i+1}/6$
Momentum	$C_t - C_{t-5}$
Psychology	$\frac{1}{6}\sum_{i=1}^{6} UD_{t-i+1}$ where $UD_i = 1$ if $C_t > C_{t-1}$, 0 if $C_t \leq C_{t-1}$
Disparity6 (6-day disparity)	$\dfrac{C_t}{MA_6} \times 100$
Disparity25 (25-day disparity)	$\dfrac{C_t}{MA_{25}} \times 100$
ROC6 (6-day price rate-of-change)	$\dfrac{C_t}{C_{t-6}} \times 100$
VR (Volume ratio)	$\dfrac{\left\{\sum_{i=1}^{6} VU_{t-i+1} + \frac{1}{2}\sum_{i=1}^{6} VS_{t-i+1}\right\}}{\left\{\sum_{i=1}^{6} VD_{t-i+1} + \sum_{i=1}^{6} VS_{t-i+1}\right\}} \times 100$ where VU_n, VD_n and VS_n mean the volume of an up, down and steady day of the stock price index, respectively, for n days
MA25 (25-day moving average)	$\frac{1}{25}\sum_{i=1}^{25} C_{t-i+1}$
ROC25(25-day price rate of change)	$\dfrac{C_t}{C_{t-25}} \times 100$

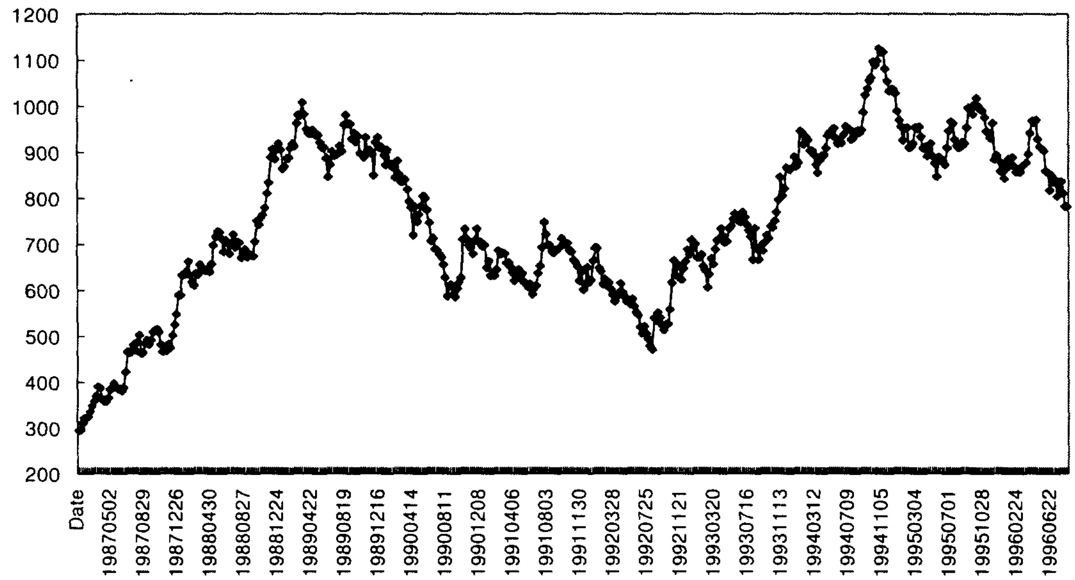

Figure 1. Weekly KOSPI data from January 1987 to August 1996

Table 2.
Models and their associated classifiers for KOSPI forecasting

Model	Classifier used in the model
Pure_TS	None
Pure_NN	None
CBR_NN	Case Based Reasoning
MDA_NN	Multivariate Discriminant Analysis
BPN_NN	Backpropagation Neural Networks

Numerical values for the performance metrics by predictive model are given in Table 3. According to RMSE, MAE and MAPE, the outcomes indicate that neural network models are superior to statistical time series model. BPN-supported neural network model and MDA-supported neural network model are superior to the pure BPN model and CBR-supported neural network model.

Our approach to integration involves a multistrategy technique which may be called second-order learning. CBR has provided the good results with the second order learning process and the integrated approach (Kim and Noh, 1997; Kim and Joo, 1997). However, CBR did not perform well in this study. In the second-order learning, the forecast from the superior method is selected on a case-by-case basis to determine the output of overall model. In other words, the second step (BPN in this study) serves as a metalevel process to determine which of three elementary modules (CBR, BPN and MDA in this study) perform better. In this point, CBR is not a good metalevel predictive method. Thus, we will choose BPN or MDA as the elementary module for real application of model.

Table 3.
Performance results in the case of KOSPI forecasting based on the root mean squared error (RMSE), the mean absolute error (MAE) and the mean absolute percentage error (MAPE)

Model	RMSE	MAE	MAPE (%)
Pure_TS	57.14	49.05	5.420
Pure_NN	28.18	22.66	2.650
CBR_NN	28.25	22.63	2.618
MDA_NN	25.71	20.52	2.377
BPN_NN	25.08	19.90	2.298

We use the pairwise t-test to examine whether there exist the differences in the predicted values of models according to the absolute percentage error (APE). This metric is chosen since it is commonly used (Carbone and Armstrong, 1982) and is highly robust (Armstrong and Collopy, 1992; Makridakis, 1993). Since the forecasts are not statistically independent and not always normally distributed, we compare the forecasts APEs using the pairwise t-test. Where sample sizes are reasonably large, this test is robust to the distribution of the data, to nonhomogeneity of variances, and to statistical dependence (Iman and Conover, 1983). Table 4 shows t-values and p-values when the prediction accuracies of the left-vertical methods are compared with those for the right-horizontal methods. Mostly, the neural network models using change-point detection perform significantly better than the time series model and the pure BPN model at a 1% significant level except the analysis of CBR-supported neural network model. Therefore, our research model is demonstrated to obtain the improved performance through the change-point detection approach.

The neural network models using change-point detection turn out to have a high potential in stock price index forecasting. This is attributable to the fact that it categorizes the input data samples into homogeneous group and extracts regularities from each homogeneous group. Therefore, the neural network models using change-point detection can cope with the noise or irregularities more efficiently than the time series model and the pure BPN model. In addition, BPN and MDA perform very well as a tool in stock price index forecasting.

Table 4.
Pairwise t-tests for the difference in residuals between the time series model, the pure BPN model and the proposed neural network models for KOSPI based on the APE with the significance level in parentheses.

Model	MDA_NN	CBR_NN	Pure_NN	Pure_TS
BPN_NN	0.885 (0.377)	2.716 (0.007)***	2.969 (0.003)***	13.151 (0.000)***
MDA_NN		2.578 (0.011)**	2.756 (0.006)***	12.498 (0.000)***
CBR_NN			0.840 (0.402)	9.846 (0.000)***
Pure_NN				9.552 (0.000)***

***Significant at 1%; **Significant at 5%

5. Concluding Remarks

This study has suggested the integrated neural network models using change-point detection in stock price index forecasting. The basic concept of the proposed model is to obtain significant intervals by change-point detection, to identify them as change-point groups, and to involve them in stock price index forecasting. We propose integrated neural network models which consist of three phases. In the first phase, we conduct the nonparametric statistical test for the change-point detection to construct the homogeneous groups. In the second phase, we apply several kinds of classifiers to forecast the change-point group. In the final phase, we apply BPN to forecast the output.

The neural network models using change-point detection perform significantly better than the time series model and the pure BPN model at a 1% significant level except the analysis of CBR-supported neural network model. Experimental results showed that the neural network models using change-point detection outperform the time series model and the pure BPN model significantly, which implies the high potential of involving the change-point detection in the model. BPN and MDA performed very well as data mining classifiers while CBR did not. Our integrated neural network models are demonstrated to be useful intelligent data analysis methods with the concept of change-point detection. In conclusion, we have shown that the proposed models improve the predictability of stock price index significantly.

The proposed model has the promising possibilities to improve the performance if further studies are to focus on the various approaches in the construction and the prediction of change-point group. In final phase of the model, other intelligent approaches can be used to forecast the final output besides BPN. In addition, the proposed models may be applied to other chaotic time series data such as exchange rate prediction. By the extension of these points, future research is expected to provide more improved neural network models with superior performances.

References

[1] Ahmadi, H. (1990). Testability of the arbitrage pricing theory by neural networks, *Proceedings of the International Conference on Neural Networks* (pp.385-393), San Diego, California.

[2] Armstrong, J.S. and Collopy, F. (1992). Error measures for generalizing about forecasting methods: Empirical comparisons, *International Journal of Forecasting,* 8, 69-80.

[3] Banergee, A., Lumsdaine, R. and Stock, J. (1992). Recursive and sequential tests of the unit root and trend break hypothesis: Theory and international evidence, *Journal of Business and Economic Statistics,* 10, 271-287.

[4] Carbone, R. and Armstrong, J.S. (1982). Evaluation of extrapolative forecasting methods: Results of academicians and practitioners, *Journal of Forecasting,* 1, 215-217.

[5] Choi, J. (1995). *Technical Indicators*. Jinritamgu Publishing, Seoul.

[6] Cochran, W.G. (1977). *Sampling Techniques,* Third edition, John Wiley & Sons, New York.

[7] Duke, L.S. and Long, J.A. (1993). *Neural network futures trading - A feasibility study,* Society for Worldwide Interbank Financial Telecomm., Adaptive intelligent systems (pp.121-132), Amsterdam: Elsevier Science Publishers.

[8] Fisher, R.A. (1936). The use of multiple measurements in taxonomic problems, *Annals of Eugenics,* 7, 179-188.

[9] Gottman, J.M. (1981). *Time-Series Analysis,* Cambridge University Press, New York.

[10] Hair, J.F. Jr., Anderson, R.E., Tatham, R.L. and Black, W.C. (1995). *Multivariate Data Analysis with Readings*, Fourth edition, Prentice Hall, New York.

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

[12] Iman, R. and Conover, W.J. (1983). *Modern Business Statistics*, Wiley, New York.

[13] Jaditz, T. and Sayers, C.L. (1995). Nonlinear in the interest rate risk premium, In R.R. Trippi (Ed), *Chaos and Nonlinear Dynamics in the Financial Markets* (pp. 335-357), Irwin, Chicago, Illinois.

[14] Kamijo, K. and Tanigawa, T. (1990). Stock price pattern recognition: A recurrent neural network approach, *Proceedings of the I. Joint Conference on Neural Networks* (pp.215-221), San Diego, California.

[15] Kaufman, K.A., Michalski, R.S. and Kerschberg, L. (1991). Mining for knowledge in databases: Goals and general description of the INLEN system, In G. Piatetsky-Shapiro and W.J. Frawley (Eds), *Knowledge Discovery in Databases* (pp. 449-462), AAAI/MIT Press, Cambridge, MA.

[16] Kim, S.H. and Joo, J. (1997). Second-order learning for complex forecasting tasks: case study of Video-On-Demand, *Korean Journal of Expert System,* 3, 31-45.

[17] Kimoto, T., Asakawa, K., Yoda, M. and Takeoka, M. (1990). Stock market prediction system with modular neural network, *Proceedings of the I. Joint Conference on Neural Networks* (pp.1-6), San Diego, California.

[18] Kohara, K., Ishikawa, T., Fukuhara, Y. and Nakamura, Y. (1997). Stock price prediction using prior knowledge and neural networks, *International Journal of Intelligent Systems in Accounting, Finance and Management*, 6, 11-22.

[19] Kolodner, J. (1991). Improving human decision making through case-based decision aiding, *AI Magazine,* 12, 52-68.

[20] Kolodner, J. (1993). *Case-Based Reasoning*, Morgan Kaufmann, San Mateo, CA.

[21] Larrain, M. (1991). Empirical tests of chaotic behavior in a nonlinear interest rate model, *Financial Analysts Journal*, Sep.-Oct., 51-62.

[22] Li, H.L. and Yu, J.R. (1999). A piecewise regression analysis with automatic change-point detection, *Intelligent Data Analysis,* 3, 75-85.

[23] Lumsdaine, R.L. and Papell, D.H. (1997). Multiple trends and the unit root hypothesis, *The Review of Economics and Staistics,* 79, 212-218.

[24] Makridakis, S. (1993). Accuracy measures: Theoretical and practical concerns, *International Journal of Forecasting*, 9, 527-529.

[25] McLachlan, G.J. (1992). *Discriminant Analysis and Statistical Pattern Recognition,* John Wiley & Sons, New York.

[26] Mishkin, F.S. (1995). The economics of money, banking, and financial markets, Harper Collins, New York.

[27] Mishkin, F.S. (1995). *The economics of money, banking, and fiancial markets,* Fourth edition, Harper Collins, New York.

[28] Oh, K.J. and Han, I. (2000). Using change-point detection to support artificial neural networks for interest rates forecasting, *Expert Systems with Applications,* 19, 105-115.

[29] Perron, P. and Vogelsang, T. (1992a). Nonstationarity and level shifts with an application to purchasing power parity, *Journal of Business and Economic Statistics*, 10, 301-320.

[30] Perron, P. (1995). Further evidence on breaking trend functions in macroeconomic variables, manuscript, Université de Montreal, Canada.

[31] Perron, P. (1990). Testing for a unit root in time series with a changing mean, *Journal of Besiness and Economic Statistics,* 8, 153-162.

[32] Perron, P. (1989). The great crash, the oil price shock, and the unit root hypothesis, *Econometrica,* 57, 1361-1402.

[33] Peters E.E. (1991). *Chaos and Order in the Capital Markets,* John Wiley & Sons, New York.

[34] Pettitt, A.N. (1979). A non-parametric approach to the change-point problem, *Applied Statistics,* 28(2), 126-135.

[35] Pettitt, A.N. (1980a). A simple cumulative sum type statistic for the change-point problem

with zero-one observations, *Biometrika,* 67, 79-84.

[36] Pettitt, A.N. (1980b). Some results on estimating a change-point using nonparametric type statistics, *Journal of Statistical Computation and Simulation,* 11, 261-272.

[37] Rapport, P. and Reichlin, L. (1989). Segmented trends and non-stationary time series, *The Economic Journal*, 99, 168-177.

[38] Rosenblatt, F. (1962). *Principles of Neurodynamics*, Spartan, New York.

[39] Rumelhart, D.E., Hinton, G.E. and Williams, R.J. (1986). Learning internal representations by back propagation, In D.E. Rumelhart, J.L. McClelland and PDP research group (Eds), *Parallel Distributed Processing*, Volume 1, MIT Press, Cambridge, MA.

[40] Trippi, R.R. and DeSieno, D. (1992). Trading equity index futures with a neural network, *Journal of Portfolio Management*, 19, 27-33.

[41] Trippi, R.R. and Turban, E. (1996). *Neural Networks in Finance and Investing*, Second edition, Irwin, Chicago.

[42] Tsaih, R., Hsu, Y. and Lai, C.C. (1998). Forecasting S&P 500 stock index futures with a hybrid AI system, *Decision Support Systems*, 23, 161-174.

[43] Vogelsang T. and Perron, P. (1995). Additional tests for a unit root allowing for a break in the trend function at an unknown time, Manuscript, Department of Economics, Ithaca, New York.

[44] White, H. (1992). Connectionist nonparametric regression: Multilayer feedforward networks can learn arbitrary mappings, In H. White (Ed), *Artificial Neural Networks: Approximations and Learning Theory*, Blackwell, Oxford, UK.

[45] Wolkenhauer, O. and Edmunds, J.M. (1997). Possibilistic testing of distribution functions for change detection, *Intelligent Data Analysis,* 1, 119-127.

[46] Yoon, Y. and Swales, G. (1991). Predicting stock price performance: A neural network approach, *Proceedings of the 24th Annual Hawaii International Conference on System Sciences* (pp.156-162), Hawaii.

[47] Zivot, E. and Andrews, D.W.K. (1992). Further evidence on the great crash, the oil-price shocks, and the unit-root hypothesis, *Journal of Business and Economic Statistics*, 10, 251-270.