Applied Mathematics & Information Sciences
An International Journal

Financial Time-Series Forecasting based on a Neural Network with Weighted Fuzzy Membership Functions and the Takagi-Sugeno Fuzzy Model

Sang-Hong Lee[1] *and Joon S. Lim*[2,*]

[1] Department of Computer Science & Engineering, Anyang University, Republic of Korea
[2] IT College, Gachon University, Republic of Korea

Abstract: This paper proposes financial time-series forecasting using a feature selection method based on the non-overlap area distribution measurement method supported in a neural network with weighted fuzzy membership functions (NEWFM) and the Takagi-Sugeno (T-S) fuzzy model. The non-overlap area distribution measurement method selects the minimum number of features with the highest performance by removing the worst features one by one. This paper uses $CPP_{n,m}$ (current price position on day n: percentage of the difference between the price on day n and the moving average of the past m days' prices from day n-1) as a technical indicator. The performance result improves from 58.35% to 58.86% when $CPP_{n,5}$ is added to a minimum number of features that are selected by the non-overlap area distribution measurement method. The T-S fuzzy model can provide the weighted average defuzzification method to represent a trend line such as a financial time series. This paper generally demonstrates fluctuations similar to a financial time series of the daily KOSPI's trend line by the defuzzification method.

Keywords: Fuzzy neural networks, feature selection, financial time series, KOSPI, technical indicator

1 Introduction

A useful technique to facilitate knowledge discovery from a database (KDD) is to select the minimum number of features with the highest performance result, thus ensuring low operation costs. Principal component analysis (PCA) is a vector space transform often used to reduce multidimensional data sets to lower dimensions for analysis. PCA arranges components in descending order according to their significance in contributing to the overall data variations [1]. Therefore, the first principal component explains most of the variation in the given data. The second principal component explains the cause of the next level of variation, and so on. However, although PCA reduces multidimensional data sets to low dimensions for analysis, it cannot reduce the number of features. Therefore, PCA does not have the general advantage of feature selection.

Existing studies of stock forecasting that are representative of time-series forecasting have mainly focused on whether stock prices will go up or down

[11][12]. However, while knowing whether stock prices will go up or down is important, identifying the trend in prices is also important. The trend of stock prices can be used as basic data in forecasting the sizes of rises and drops.
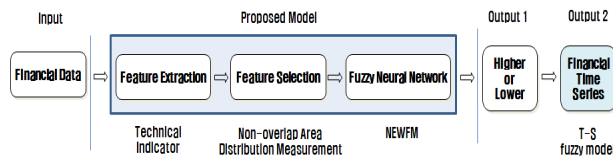
This paper suggests financial time-series forecasting using a feature selection method based on the non-overlap area distribution measurement method supported in a neural network with weighted fuzzy membership functions (NEWFM) [7][10] and the Takagi-Sugeno (T-S) fuzzy model [8]. The non-overlap area distribution measurement method removes the worst features one by one, and then selects the minimum number of features, each of which constructs an interpretable fuzzy membership function. All features are interpretably formed in weighted fuzzy membership functions, preserving the disjunctive fuzzy information and characteristics. The NEWFM's performance results are evaluated by the same number of features used by Kim [2], and also by the minimum number of features selected by the non-overlap area distribution measurement

* Corresponding author e-mail: jslim@gachon.ac.kr

method. We propose a financial time-series forecasting model that is represented by the weighted average defuzzification method of the T-S fuzzy model.

2 Overview of Financial Time-Series Forecasting Model

Figure 2.1 shows the financial time-series forecasting model, which consists of four steps. In the first step, 12 technical indicators are selected to preprocess the financial data. In the second step, the minimum number of features is selected by the non-overlap area distribution measurement method based on a NEWFM [7][10]. In the third step, the NEWFM forecasts the highest performance results with the minimum number of features. In the final step, the T-S fuzzy model is used to support the financial time series.

Figure 2.1: Diagram of financial time-series forecasting model.

2.1 Experimental data

This paper used 2,928 trading days, from January 1989 to December 1998, which was the total number of samples used by Kim, and also the 12 technical indicators selected by Kim [2], to forecast changes in the daily Korean composite stock price index (KOSPI). Kim divided the samples into two subsets, training sets and holdout sets, which included 2,347 and 581 trading days, respectively. Table 1 shows the 12 technical indicators and their formulas [2]. This paper aimed to forecast changes in the daily KOSPI. Increases and decreases in the KOSPI were classified as "1" and "2" respectively; "1" means that the next day's data are lower than today's data, and "2" means that the next day's data are higher than today's data.

2.2 Neural network with weighted fuzzy membership function (NEWFM)

A NEWFM [7][10][13] is a supervised classification neuro-fuzzy system using the bounded sum of weighted fuzzy membership functions (BSWFMs). The previous structure of the NEWFM comprises three layers: the input, hyperbox, and class layers. The new structure of the NEWFM, illustrated in Figure 2.2, comprises four layers: the input, hyperbox, class, and output layers. The input layer contains n input nodes for an n-featured input

pattern. The hyperbox layer consists of m hyperbox nodes. Each hyperbox node B_l to be connected to a class node contains n BSWFMs for n input nodes. The class layer is composed of p class nodes, each of which is connected to one or more hyperbox nodes. In this paper, the output layer is designed to support financial time-series forecasting based on the T-S fuzzy model [8]. The hth input pattern can be recorded as $I_h = \{A_h = (a_1, a_2, ..., a_n), class\}$, where $class$ is the result of classification and A_h is n features of an input pattern.
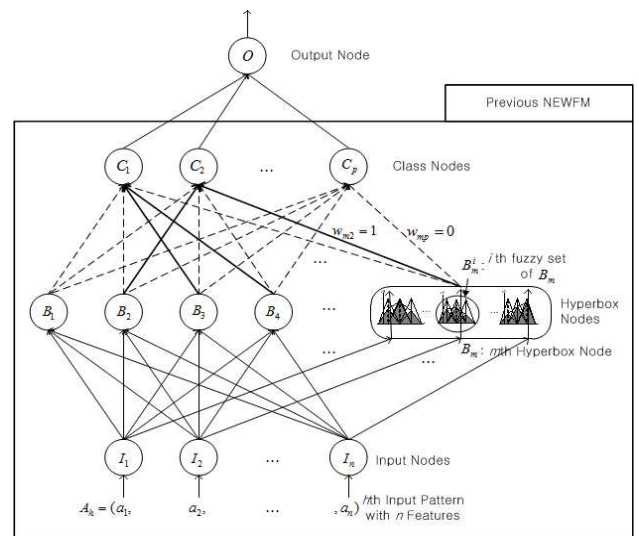
Figure 2.2: Structure of the NEWFM.

3 Feature Selection

Selecting the number of fuzzy rules and identifying the important features have received attention in recent literature [3][4][6][9]. In this paper, 4 minimum features were selected by the non-overlap area distribution measurement method [7][10] from 12 initial features. The method measures the degree of salience of the ith feature by non-overlapped areas with the area distribution by means of the following equation [5]:

$$f(i) = (Area_H^i + Area_L^i)^2 / \frac{1}{(1 + e^{-|Area_H^i - Area_L^i|})} \quad (1)$$

where $Area_H$ and $Area_L$ are the higher and lower phase superior areas, respectively. As an example, for the good and bad candidate features for feature selection among 12 initial features, $Area_H$ and $Area_L$ are shown in Figure 3.1. The larger the value of $f(i)$, the more the feature's characteristic is implied.

Table 1. Initial features and their formulas

Feature name	Formula
Stochastic %K	$\frac{C_t - LL_{t-5}}{HH_{t-5} - LL_{t-5}} \times 100$
Stochastic %D	$\frac{\sum_{i=0}^{n-1} \%K_{t-i}}{n}$
Stochastic Slow %D	$\frac{\sum_{i=0}^{n-1} \%D_{t-i}}{n}$
Momentum	$C_t - C_{t-4}$
ROC (Rate of Change)	$\frac{C_t}{C_{t-n}} \times 100$
Larry William's %R	$\frac{H_n - C_t}{H_n - L_n} \times 100$
Accumulation /Distribution (A/D) Oscillator	$\frac{H_t - C_{t-1}}{H_t - L_t}$
Disparity 5 days	$\frac{C_t}{MA_5} \times 100$
Disparity 10 days	$\frac{C_t}{MA_{10}} \times 100$
Price Oscillator (OSCP)	$\frac{MA_5 - MA_{10}}{MA_5}$
Commodity Channel Index (CCI)	$\frac{M_t - SM_t}{0.015 \times D_t}$
Relative Strength Index (RSI)	$100 - \frac{100}{1 + (\sum_{i=0}^{n-1} Up_{t-i}/n)/(\sum_{i=0}^{n-1} Dw_{t-i}/n)}$

C_t is the closing price at time t, L_t is the low price at time t, LL_t is the lowest low price in the past t days, H_t is the high price at time t, HH_t is the highest high price in the past t days, MA_t is the moving average of t days, M_t is $(H_t + L_t + C_t)/3$, SM_t is $\sum_{i=1}^{n} M_{t-i+1}/n$, D_t is $\sum_{i=1}^{n} |M_{t-i+1} - SM_t|/n$, Up_t is an upward price change, and Dw_t is a downward price change.
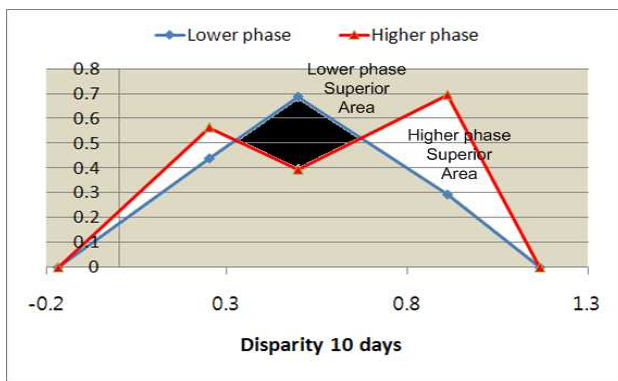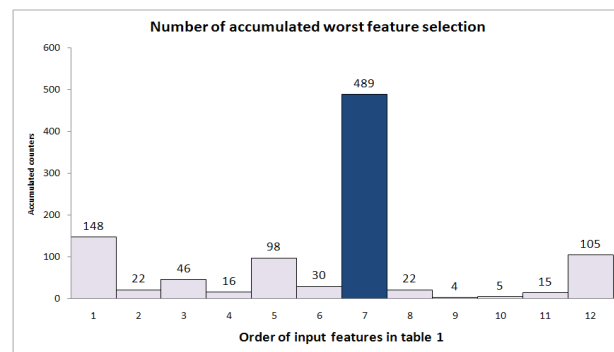

(a) An example of a good candidate feature for feature selection.
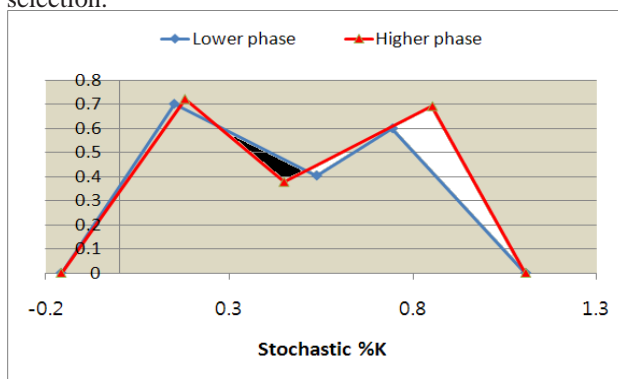
(b) An example of a bad candidate feature for feature selection.

Figure 3.1: $Area_H$ (white) and $Area_L$ (black) for the good and bad candidate features for feature selection among 12 initial features.

Figure 3.2: Numbers of accumulated features with the smallest non-overlap area for feature selection.

The worst of the 12 initial features are removed one by one by the non-overlap area distribution measurement method [10]. The training processes and testing processes were iterated 1000 times. Figure 3.2 shows a state where the minimum formula $f(i)$ values were accumulated for all features during 1000 iterations of the experiment. The 7th feature was judged to be the worst one, and hence, the 7th feature was deleted before the next experiment started, and then 4 minimum features with the highest performance results are finally selected for the holdout sets. In Table 2, "√" refers to the features selected from the 12 initial features.

This experiment created two hyperboxes for classification. A hyperbox that contains a set of lines (BSWFM) in Figure 3.1 is a rule for class 1 (the lower phase), and the other hyperbox that contains a set of lines (BSWFM) is another rule for class 2 (the higher phase). The graphs in Figure 3.1 were obtained from the NEWFM program's training process, and graphically show the difference between the lower and higher phases for each feature. The lower phase means that the next

day's data are lower than today's data. The higher phase means that the next day's data are higher than today's data.

4 Experimental Results

4.1 Performance results

Kim compared support vector machines (SVMs) with back-propagation neural networks (BP) for forecasting changes in the daily KOSPI [2]. The performance results of the NEWFM shown in Table 3 were evaluated by 12 initial features, which are presented in Table 1. As Table 3 shows, the SVM outperformed the NEWFM by 1.03% for the holdout data in the experiment that used 12 initial features. However, the NEWFM outperformed the SVM by 0.52% for the holdout data in the experiment that used 4 minimum features selected by the non-overlap area distribution measurement method.

Table 2. Description of the selected features

Feature name	Selected
Stochastic %K	
Stochastic %D	√
Stochastic Slow %D	
Momentum	√
ROC (Rate of Change)	
Larry William's %R	
Accumulation /Distribution (A/D) Oscillator	
Disparity 5 days	√
Disparity 10 days	√
Price Oscillator (OSCP)	
Commodity Channel Index (CCI)	
Relative Strength Index (RSI)	

Table 3. Comparisons of performance results for Kim with NEWFM using 12 initial features

	NEWFM	SVM	BP
Performance results	56.80%	57.83%	54.73%

Table 4. Comparisons of performance results for Kim with NEWFM using 4 minimum features

	NEWFM	SVM	BP
Performance results	58.35%	57.83%	54.73%

4.2 Improving performance results using our own technical indicator

This paper used $CPP_{n,m}$ (current price position on day n: percentage of the difference between the price on day n and the moving average of the past m days' prices from day n-1) as a technical indicator to forecast changes in the daily KOSPI [10]. $CPP_{n,m}$ is calculated by [10]:

$$CPP_{n,m} = ((C_n - MA_{n-1,n-m})/MA_{n-1,n-m}) \times 100 \quad (2)$$

where C_n is the closing price on day n and $MA_{n-1,n-m}$ is the moving average price of the past m days from day n-1. Table 5 shows that the performance result increased from 58.35% to 58.86% when $CPP_{n,5}$ was added to 4 minimum features that were selected by the non-overlap area distribution measurement method.

Table 5. Comparison of performance results for NEWFM using 5 features to NEWFM using 4 features

|	NEWFM using 5 features	NEWFM using 4 features
Performance results	58.86%	58.35%

4.3 Financial time series based on the Takagi-Sugeno (T-S) fuzzy model

The T-S fuzzy model [8], a powerful tool for modeling complex nonlinear systems, has consequent parts consisting of linear functions and can be viewed as an expansion of piecewise linear partition.

$$R^i : IF\, x_1\, is A_1^i, ..., x_m\, is A_m^i\, THEN y^i = a_0^i + ... + a_m^i x_m \quad (3)$$

$$y = \frac{\sum_{i=1}^c w^i y^i}{\sum_{i=1}^c w^i}\, where\, w^i = Min\{A_1^i(x_1), ..., A_m^i(x_m)\} \quad (4)$$

where $R^i(i = 1, 2, ..., c)$ denotes the ith fuzzy rule, $x_i(i = 1, 2, ..., m)$ are the input variables, y^i are the rule output variables, $A_1^i, ..., A_m^i$ are fuzzy sets of the ith rule for x_i, and $a_0^i, ..., a_m^i$ are the parameter sets in the consequent part.
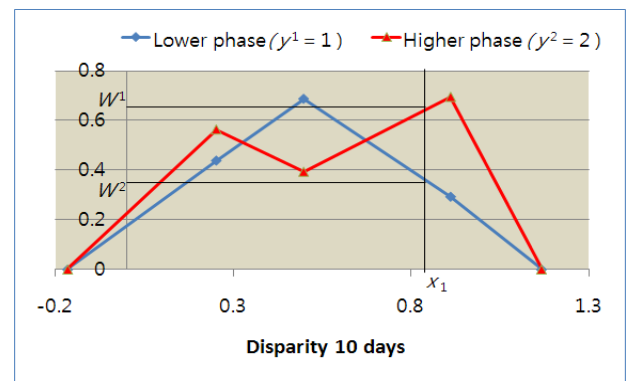

**Figure 4.1:** An example of how to calculate defuzzification value.

Figure 4.1 shows examples of values to be used as inputs of Eq. (4). Figure 4.2 shows that although both the two stock prices increase, the sizes of their rises are different. Using the defuzzification values of the T-S fuzzy model, the sizes of stock price rises can be shown as per Figure 4.2. In addition, by successively connecting the defuzzification values in Figure 4.2, a trend line such

as a financial time series can be represented as shown in Figure 4.3. The financial time-series forecasting of NEWFM can be represented by the weighted average defuzzification method of the T-S fuzzy model. Figure 4.3 shows the financial time series of the daily KOSPI. This result generally demonstrates fluctuations similar to those of the daily KOSPI's trend line.
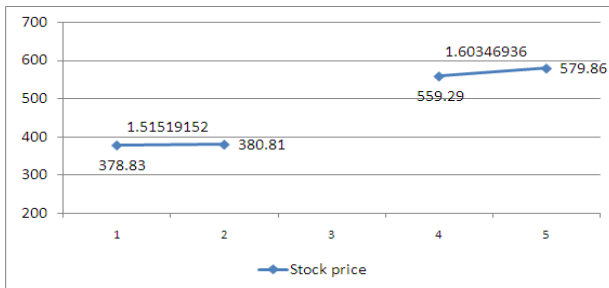
Figure 4.2: An example of defuzzification value on stock's increases.
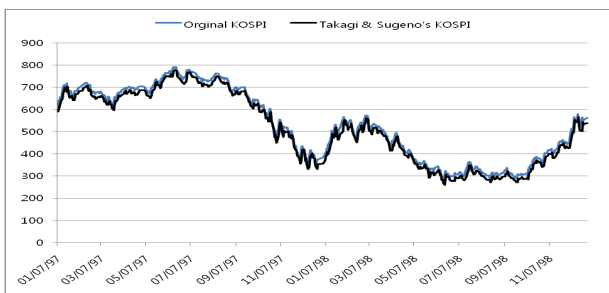
Figure 4.3: Comparison of the original daily KOSPI and the weighted average defuzzification method of the T-S fuzzy model.

5 Conclusion

This paper proposes financial time-series forecasting using a feature selection method based on the non-overlap area distribution measurement method, which is based on a neural network with weighted fuzzy membership functions (NEWFM), and on the weighted average defuzzification method of the T-S fuzzy model. In this paper, the non-overlap area distribution measurement method selected 4 minimum features with the highest performance result from 12 initial features by removing the worst features one by one. The $CPP_{n,m}$ (current price position on day n: percentage of the difference between the price on day n and the moving average of the past m days' prices from day n-1) was used as a technical indicator. The performance result improves from 58.35% to 58.86% when the $CPP_{n,5}$ is added to the 4 minimum features that were selected by the non-overlap area distribution measurement method as new. The financial time-series forecasting of an NEWFM can be represented by the weighted average defuzzification method of the T-S fuzzy model.

Acknowledgement

References

[1] Sanger T. D., Optimal Unsupervised Learning in a Single-Layer Linear Feedforward Neural Network, Neural Networks, **12**, 459-473 (1989).

[2] Kim K.-J., Financial time series forecasting using support vector machines, Neurocomputing, **55**, 307-309 (2003).

[3] Lutu P. E. N. and Engelbrecht A. P., A decision rule-based method for feature selection in predictive data mining, Expert Systems with Applications, **37**, 602-609 (2010).

[4] Zhou S.-M and Gan J.Q., Constructing L2-SVM-Based Fuzzy Classifiers in High-Dimensional Space With Automatic Model Selection and Fuzzy Rule Ranking, IEEE Trans. on Fuzzy Systems, **15**, 398-409 (2007).

[5] Lee S.-H. and Lim J. S., Minimized Stock Forecasting Features Selection by Automatic Feature Extraction Method, Korean Institute of Intelligent Systems, **19**, 206-211 (2009).

[6] Setnes M. and Roubos H., GA-Fuzzy Modeling and Classification: Complexity and Performance, IEEE Trans. Fuzzy Systems, **8**, 509-522 (2000).

[7] Lim J. S., Finding Features for Real-Time Premature Ventricular Contraction Detection Using a Fuzzy Neural Network System, IEEE TRANSACTIONS ON NEURAL NETWORKS, **20**, 522-527 (2009).

[8] Takagi T. and Sugeno M., Fuzzy identification of system and its applications to modeling and control, IEEE Trans. Syst., Man, Cybern., **15**, 116-132 (1985).

[9] Sotoca J. M. and Pla F., Supervised feature selection by clustering using condi-tional mutual information-baseddistances, Pattern Recognition, **43**, , 2068-2081 (2010).

[10] Lee S.-H. and Lim J. S., Forecasting KOSPI based on a neural network with weighted fuzzy membership functions, Expert Systems with Applications, **38**, 4259-4263 (2011).

[11] Kim K.-J., Artificial neural networks with evolutionary instance selection for financial forecasting, Expert Systems with Applications, **30**, 519-526 (2006).

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

[13] Lee S.-H. and Lim J. S., Comparison of DBS and levodopa on resting tremor using a fuzzy neural network system, Measurement, **46**, 1995-2002 (2013).

Sang-Hong Lee received the B.S., M.S., and Ph.D. degrees in computer science from Kyungwon University, Korea in 1999, 2001, and 2012, respectively. He is currently an assistant professor in the department of computer science & engineering at Anyang University, Korea. His research focuses on neuro-fuzzy systems, stocks prediction systems, and biomedical prediction systems.

Joon S. Lim received his B.S. and M.S. degrees in computer science from Inha University, Korea, the University of Alabama at Birmingham, and Ph.D. degree was from Louisiana State University, Baton Rouge, Louisiana, in 1986, 1989, and 1994, respectively. He is currently a professor in the department of computer software at Gachon University, Korea. His research focuses on neuro-fuzzy systems, biomedical prediction systems, and human-centered systems. He has authored three textbooks on Artificial Intelligence Programming (Green Press, 2000), Javaquest (Green Press, 2003), and C# Quest (Green Press, 2006).