國立政治大學理學院應用物理研究所

碩士論文

Graduate Institute of Applied Physics, College of Science

National Chengchi University

Master Thesis

基於 EEMD 之類神經網路預測方法進行台指選擇權交易策略

TAIEX Option Trading by using EEMD-based Neural Network Learning Paradigm

李恩慈

Li En Tzu

指導教授：蕭又新博士、廖四郎博士

Advisor：Shiau Yuo Hsien, Liao Szu Lang

中華民國一〇〇年六月

June 2011

Contents

Tables

Figures

摘　要

金融市場瞬息萬變，幾乎所有商品價格都是非線性的動態過程，如何預測價格一直都是倍受討論和研究的議題。隨著電腦科技的不斷進步,許多財務學者以市場上的歷史交易資料作為研究對象，希望能夠預測出有效的結果。本研究利用 EEMD 法拆解原始加權指數訊號，建立類神經網路模型，並預測出未來市場之價格後，利用 FK 值當作交易門檻，帶回台指選擇權做交易測試並計算報酬。由於不同神經元個數會配適出不同的預測結果，本研究希望能夠找到較適合使用在指數預測的網路架構。

關鍵字：EEMD、ANN、預測模型、交易策略、FK 指標

Abstract

The financial market forecasting is characterized by data intensity, noise, non-stationary, high degree of uncertainty, and hidden relationships. Investors are concerned about the forecasting market price. Throughout the development of computational technology, researchers have been involved in data mining on historical trading enabling them to have a more accurate data. This research uses Ensemble Empirical Mode Decomposition-based Artificial Neural Networks (ANNs) learning paradigm to provide different ways to analyze the stock market. In our research, we used the ANN method to obtain our prediction of the stock price. First, the previous day's stock price needs to be decomposed in order to see the various variables, that is, the numerous IMFs seen on the graphs. Acquiring the information, it is inserted into the ANN method to get a prediction. Following that, the prediction can then be transformed into a simpler result via the Forward Calculator % K indicator. As a result, the FK value can display a signal if to buy or sell, and confirm trading time, and make buy or sell Call-Put decisions on TAIEX options. In summary, we found different neuron numbers in the hidden layers that may affect the result of prediction.

Keywords: *Ensemble Empirical Mode Decomposition, Artificial Neural Network, Forecasting, and FK Indicator.*

1. Introduction

In the past, there have been various studies concerning the forecasting of the stock market's issues. They have become one of the major areas of research in financial market. The financial market is a complex, evolutionary, and non-linear dynamical system (Abu-mostafa and Atiya, 1996). The financial market forecasting is characterized by data intensity, noise, non-stationary, high degree of uncertainty, and hidden relationships. Quantitative methods are crucial for forecasting purposes in the stock market as well as for improved decisions and investments. Predicting stock data with traditional time series analysis has proven to be difficult. Artificial Neural Networks (ANNs) provide different ways to analyze the stock market. Primarily, there has not been any assumption of a suitable mathematical model to be made prior to forecasting. Furthermore, a neural network has the ability to extract useful information from large sets of data, which often is required for a satisfying description of a financial time series. This thesis focuses on the development of a stock market forecasting model based on artificial neural network architecture. This study constructs EEMD-based ANN algorithm as a baseline model, which enables us to test the TAIEX option return.

With ANNs, there is no need to specify a particular model form. Rather, the model is adaptively formed based on the features presented in the data. Chan et al. (2000) used daily trade data of the listed companies from Shanghai Stock Exchange to collect technical analysis with the means of neural networks. Two learning algorithms and two weight initializations were compared. The results demonstrated that neural networks can model the time series satisfactorily, whichever learning algorithm and weight initialization are adopted. Hamid et al. (2004) discussed the ability to foretell the volatility of the markets is critical to analysts. Among the large array of approaches available for forecasting volatility, neural networks are gaining in popularity. They present a primer for using neural networks for financial forecasting. They differentiate between volatility forecasts and neural networks with implied volatility from S&P 500 Index futures options using the Barone-Adesi and Whaley (BAW) American futures options pricing model to price a model. Forecasts from neural networks' outperformance implied volatility forecasts and were not found to be significantly different from realized volatility.

Lin and Yu (2009) transformed ANN prediction into a simple FK indicator for decision making trading strategy, whose profitability is evaluated against a simple buy-hold strategy. They adopted the neural network approach to analyze the Taiwan Weighted Index and the S&P 500 in the States. Consequently, they–found that the trading rule based on ANNs generated higher returns than the buy-hold strategy.

Huang et al. (1998) introduced an Empirical Mode Decomposition (EMD) method to be more efficient in the prediction of the economic dilemma. EMD decomposes the original data into a small number of independent and nearly periodic intrinsic modes based on a local characteristic scale. Since data is the only link we have with the reality and by exploring data's intrinsic modes, EMD not only helps discover the characteristics of the data but also helps understand the underlying rules of reality. Huang et al. (2003) used a weekly mean of the thirty-year mortgage rate covering the period from January 1972 to December 2000 in the EMD method. As most of the financial data are inherently non-stationary and non-linear, it is important that they adopt a method designed for such processes. The insouciant assumptions of stationary and homogeneous steps and many other similar ones need to be carefully scrutinized. They feel that this new method deserves a trial in this new area of financial data analysis.

Wu and Huang (2004) improved the EMD method by sifting an ensemble of white noise-added signal and treating the mean as the final true result. The new method called Ensemble Empirical Mode Decomposition (EEMD) is finite and not infinitesimal. Amplitude white noise is necessary to force the ensemble to exhaust all possible solutions in the sifting process, thus making the different scale signals to collate into the proper intrinsic mode functions (IMFs) dictated by the dyadic filter banks. As EEMD is a time–space analysis method, the added white noise is averaged out with sufficient number of trials; the only persistent part that survives the averaging process is the component of the original data, which is then treated as the true and more physical meaningful answer. The effect of the added white noise will result in a uniform reference frame in the time–frequency space; therefore, the added noise collates the portion of the signal displayed on the comparable scale in one IMF. With this ensemble mean, one can separate scales naturally without any prior subjective criterion selection in the intermittence test for the original EMD algorithm.

This new approach utilizes the full advantage of the statistical characteristics of white noise to perturb the signal in its true solution neighborhood, and to cancel itself out after serving its purpose; therefore, it represents a substantial improvement over the original EMD and is a truly noise-assisted data analysis (NADA) method.

Klevecka and Lelis (2008) (created a functional algorithm of preprocessing of input data taking into account the specific aspects of teletraffic and properties of neural networks.) kind of confusing The practical application to forecasting telecommunication data sequences shows that the procedure of data preprocessing decreases the time of learning and increases the plausibility and accuracy of the forecasts. The algorithm of pre-processing input data neural networks was applied to several real time series of different lengths representing the intensity of telephone traffic and the intensity of the total and international outgoing traffic of the IP network. They found one of the obvious advantages of neural networks is that they can work successfully with non-normally distributed data.

Zhang et al. (2008) used EEMD method to explain the generation of time series data from a novel perspective. They concern the importance of understanding the underlying characteristics of international crude oil price movements, which attracts much attention from academic researchers and business practitioners. Due to the intrinsic complexity of the oil market, however, most of them failed to produce consistently good results. With different time ranges and frequencies are decomposed into several independent intrinsic modes, from high to low frequency. The intrinsic modes are composed into a fluctuating process, a slowly varying part and a trend based on fine-to-coarse reconstruction. The economic meanings of the three components are identified as short term fluctuations caused by normal supply-demand disequilibrium or some other market activities, the effect of a shock of a significant event, and a long-term trend. They claim EEMD is shown to be a vital technique for crude oil price analysis.

Yu et al. (2008) used an EMD based neural network ensemble-learning paradigm which was proposed for the world crude oil spot price forecasting. For this purpose, the original crude oil spot price series were first decomposed into a finite, and often small, number of intrinsic mode functions (IMFs). Then a three-layered feed- forward neural network (FNN) model was used to model each of the extracted

IMFs, so that the tendencies of these IMFs could be accurately predicted. Finally, the prediction results of all IMFs are combined with an adaptive linear neural network (ALNN), to formulate an ensemble output for the original crude oil price series. For verification and testing, two main crude oil price series, West Texas Intermediate (WTI) crude oil spot price and Brent crude oil spot price, are used to test the effectiveness of the proposed EMD-based neural network ensemble learning methodology. Empirical results obtained demonstrate attractiveness of the proposed EMD-based neural network ensemble-learning paradigm. This study proposes using an EMD-based neural network ensemble-learning model to predict world crude oil spot prices. In terms of empirical results, we find that across different forecasting models, for the two main crude oil prices, WTI crude oil spot price and Brent crude oil spot price, in terms of different criteria, the EMD-based neural network ensemble-learning model performs the best. In all testing cases, the RMSE is the lowest, indicating that the EMD-based neural network ensemble forecasting paradigm can be used as a very promising methodology for world crude oil price prediction.

Yu et al. (2010) provide a multiscale neural network learning paradigm to predict financial crisis events for early-warning purposes. In the proposed multiscale neural network learning paradigm, currency exchange rate, a typical financial indicator that usually reflects economic fluctuations, is first chosen. Using the neural network weights, some important intrinsic mode components (IMCs) are selected as the final neural network inputs and some unimportant IMCs that are of little use in mapping from inputs to output are discarded. Using these selected IMCs, a neural network learning paradigm is used to predict future financial crisis events, based upon some historical data. Experimental results reveal that the proposed multiscale neural network learning paradigm can significantly improve the generalization performance relative to conventional neural networks.

Djennas et al. (2011) presents a hybrid model for predicting the occurrence of currency crises by using the Artificial Intelligence tools. The model integrates the learning ability of the Artificial Neural network (ANN) with the inference mechanism of Empirical Mode Decomposition (EMD). Thus, for better detection of currency crises emergence, an EMD-ANN model based on event analysis approach is proposed. The time series to be analyzed are first decomposed into several intrinsic modes with different time scales. Consequently, the different intrinsic modes are then explored by a neural network model in order to predict future crisis financial events. For illustration purposes, the proposed EMD-ANN learning approach is applied to exchange rate data of Turkish Lira to evaluate the state of financial crisis. The empirical results show that the proposed EMD-ANN model leads to a good prediction of crisis. Significantly, the proposed model can thus lead to a somewhat more prescriptive modeling approach based on the determination of causal mechanisms towards finding ways to prevent currency crises.

In this research, we want to modify a better decomposition method by using EEMD and bring all IMFs into ANN as neurons to forecasting stock market price. The next section is methodology of this paper. Section 3 describes the index option in Taiwan; Section 4 examines performance patterns of our experiment period. Section 5 concludes.

2. Methodology

2.1 The Ensemble Empirical Mode Decomposition (EEMD)

Empirical Mode Decomposition (EMD)

Compared to traditional analysis method, the empirical mode decomposition (EMD, Huang et al., 1998) is intuitive, direct, and adaptive, with a posterior-defined basis, from the decomposition method, based on and derived from the non-linear and non-stationary data. The decomposition has an assumption that any data consists of different simple intrinsic mode of oscillations. Each intrinsic mode, no matter linear or not, represents an oscillation called IMF. The condition of IMF will have the same number of extreme and zero-crossings, and will be symmetric with respect to the local mean. Usually, the data may have many different oscillations; those can be represented by the intrinsic mode functions (IMFs) with following sifting process.

Sifting process

1) Identify local extremes in the experimental respiratory time series data x_t
2) Link all local maxima and minima by cubic spline interpolation to generate its upper and lower envelops $e_{t,\max}$ and $e_{t,\min}$
3) Calculate mean value $m_{t,1}$ from envelops as

$$m_{t,1} = \frac{e_{t,\max} + e_{t,\min}}{2} \qquad (2.1)$$

4) Calculate the difference between x_t and $m_{t,1}$ then we get

$$h_{t,1} = x_t - m_{t,1} \qquad (2.2)$$

5) Check the properties of $\mathrm{h}_{t,1}$, if it doesn't satisfy the condition, then repeat step (1) to step (4) Until $h_{t,k}$ satisfies the stopping criterion SD:

$$h_{t,1} = x_t - m_{t,1}$$

$$h_{t,2} = h_{t,1} - m_{t,2}$$

$$\vdots$$

$$h_{t,k} = h_{t,k-1} - m_{t,k} \;;\; 0.2 \le SD = \sum_{t=1}^{T} \frac{(h_{t,k-1} - h_{t,k})^2}{h_{t,k-1}^2} \le 0.3 \qquad (2.3)$$

6)Then we set first component $h_{t,k}$ as first IMF $c_{t,1}$

7) Set residue $r_{t,1}$ then repeat step (1) to (6) n times to separate the following IMFs and residues from original data.

$$x_t - c_{t,1} = r_{t,1}$$

$$r_{t,1} - c_{t,2} = r_{t,2}$$

$$\vdots$$

$$r_{t,n-1} - c_{t,n} = r_{t,n} \qquad (2.4)$$

The sifting process stops when these predetermined criteria:

a) Either when the component $c_{t,n}$ or the residue $r_{t,n}$ becomes so small that it is less than the predetermined value of substantial consequence.

b) The final residue $r_{t,n}$ will be a constant or a monotonic function that represents the general trend of the time series.

8) At the end of the sifting process, the original time series can be expressed as

$$x_t = \sum_{i=1}^{n} c_{t,i} + r_{t,n} \qquad (2.5)$$

Where n is the number of IMFs; $r_{t,n}$ is the final residue, also the trend of x_t and $c_{t,i}$ represents IMFs that are nearly orthogonal to each other. After the sifting process, the original data set is decomposed into these IMFs which is represented from high frequency to low frequency, and every IMFs may have its own physical meaning.

Ensemble Empirical Mode Decomposition (EEMD)

EMD is proved to be a useful data analysis method for extracting signals from nonlinear and non-stationary data. However, EMD still has its defect, a single IMF may either consist of signals in widely disparate scales, or a signal of similar scale belonging to different IMF components called mode mixing. Thus, Wu and Huang (2004) modified Ensemble Empirical Mode Decomposition (EEMD) to overcome the problem, which provides a method to determine whether an IMF contains true signals based on statistical comparison with the white noise. In this paper, we used an EEMD methodology to separate out the embedded frequencies in economic variables. EEMD came about because of the various problems with EMD associated with 'mode mixing'. 'Mode mixing' occurs when either different frequencies are found to reside in the same IMF or when similar frequencies are found to reside in different IMFs.

We know that each observed data were mixed with true time series and noise, and even if the data were collected by individual observations with different noise levels, the ensemble mean will still be close to the true time-series. This means that we can extract the true signal from data by adding some white noise. The followings are the EEMD processes:

1) Add a white noise series to the original data

$$y_t^i = x_t^i + \omega_t^i \tag{2.6}$$

x_t^i is observation data, ω_t^i is added white noise

2) Decompose the data with added white noise into IMFs
3) Repeat the above two steps iteratively, and add different white noise each time; finally we obtain the ensemble means of corresponding decompositions' IMFs

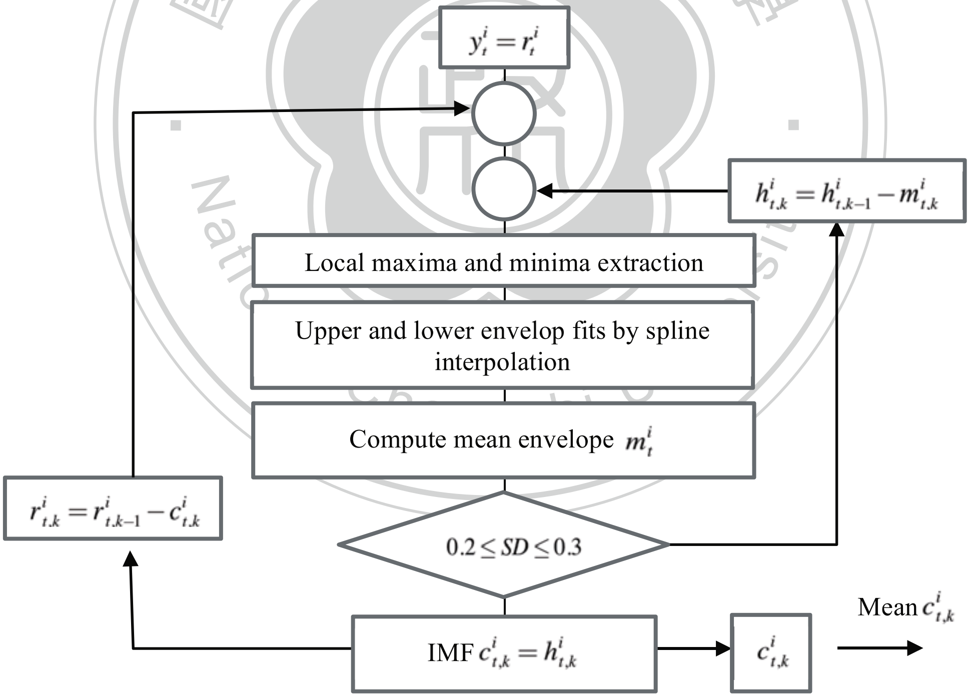

Figure 2.1 Flowchart of EEMD Sifting Process

However, there is a well-established statistical rule proven by Wu and Huang (2004) to control the effect of adding white noise:

$$\varepsilon_n = \frac{\varepsilon}{\sqrt{N}} \tag{2.7}$$

Which N is the number of ensemble members, ε is the amplitude of the added noise and ε_n is the final standard deviation of error defined as the difference between the input signal and the corresponding IMFs. Empirically, the number of the ensemble members N is always set to 100 and the ε_n always set to 0.1 or 0.2 (Zhang et al., 2008). The procedure of adding white noise successfully makes signals of comparable scales to collate into one IMF. Therefore, the EEMD successfully reduces the occurrence of mode mixing and is a substantial improvement over the original EMD.

In this study, we simulated 3 different values of ε (equal to 0.1, 0.2 and 0.3) to discuss if adding different white noise levels may effect the decomposed original signal. Following that, the EEMD decompositions of TAIEX with noise was added to see any affects on the signals. Red, green, and black lines correspond to EEMD decompositions with added noise of standard deviation of 0.1, 0.2, and 0.3, respectively. The ensemble number for each case is 100.

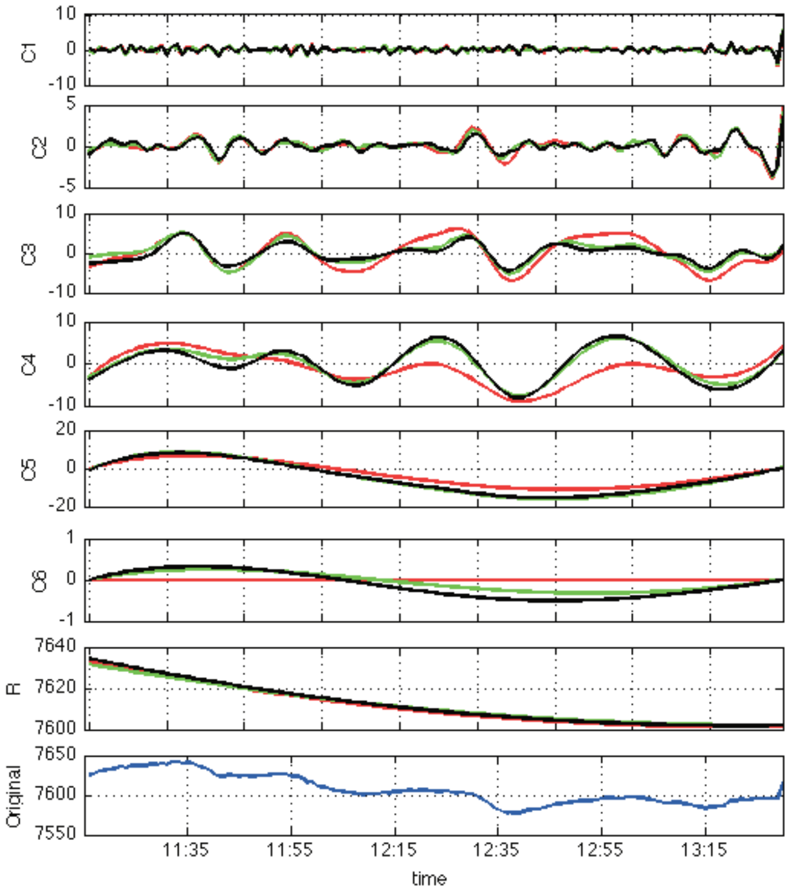

Figure 2.2 EEMD Decompositions of TAIEX with Different Noise Level

All the IMFs and residue data for the TAIEX are from Aug. 31 between 11:16 AM to 1:30 PM .As a result, we can find residue means' direction in this period of time, which can provide a more accurate and comprehensive reflection of each trading day or within a certain period of strength between long and short sides. Within the given period of time, the alteration of long and short sides measure the ability of bulls and bears setting a trend on graph. In a day or a period of time, the advantage of long and short sides continue to alternate with the two sides are likely to dominate in a given period. If investors observe the stock market to be in a bullish state, then the stock market will continue to rise. If investors think the market is meeting the bearish advantage, then this indicates that the share price is falling. Then the long sides will have the general balance of power, and the share price will fluctuate in a narrow region.

2.2 The Artificial Neural Networks (ANNs)

The Artificial Neural Networks (ANNs) are flexible computing frameworks for modeling a broad range of nonlinear problems. They are widely used in some practical application domains. ANN can be described as a type of multiple regressions with neurons arranged into layers and the connected neurons classified in different layers according to its connection strength called weights. A neural network can be viewed as neurons that are arranged in layers. A neuron is a processing element that processes the received input to generate an output. It is arranged in a three-layer feed forward neural network. Each layer is called according to the following: the input layer presents data into the network. Hidden layer creates an internal mapping of the input data. The Output layer receives data from the hidden layer. Each input value in the neurons are connected to each layer via different weights.

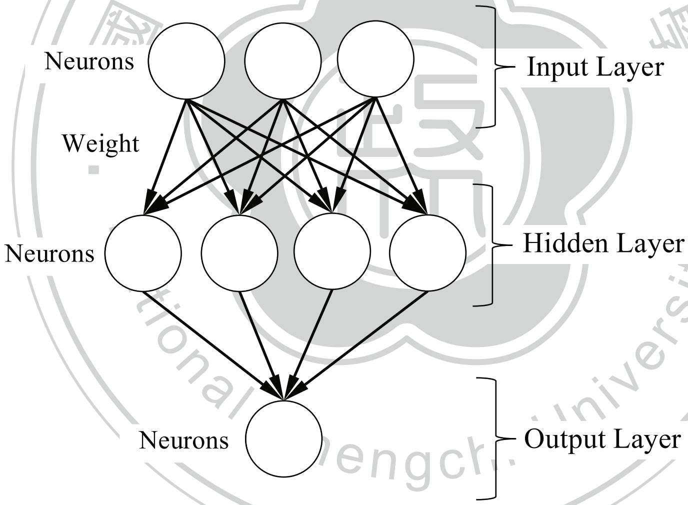

Figure 2.3 Three-Layer Feed-forward Neural Networks

Network parameters vary in 3 factors:

1) The number of hidden layers
2) The number of hidden neurons
3) Transfer function

The hidden layers provide the network with its ability to generalize. There are no standard rules available for determining the appropriate number of hidden layers would render a better generalization. In practice, the neural networks with one and occasionally two hidden layers are widely used and have performed very well. Increasing the number of hidden layers also increases computing time and over-fitting leads that might limit forecasting performance. Over-fitting occurs when the forecasting model has relatively few observations in relation to its parameters and therefore it has the ability to memorize individual points rather than learn the general pattern (I. Kaastra, M. Boyd, 1996).

The most favorable numbers were used as hidden neurons in this experimentation. Normally, it is determined through the trial route first. If there are too few of neurons inserted, this will limit the network from correctly mapping the input and output. However, if there are too many neurons entered, it may cause the network to memorize trivial patterns that can weaken its ability to demonstrate expected features or trends and make an appropriate generalization. The number of neurons can still range from one-half to two times (Mendelsohn, 1993) the geometric pyramid rule value depending on the complexity of the problem. In this study, we kept all the other parameters constant to test the numbers of hidden neurons that would perform the best.

Transfer functions are mathematical formulas that determine the output of processing neurons. Also, the purpose of it is to prevent outputs from reaching very large values. The logistic function is often used in the hidden layer transfer function, which is used to gather the units from the different layers. They map out the neurons that were entered to receive an output where the neurons are the links between the layers. The widely used transfer functions are the *Hyperbolic Function* and the *Sigmoid Function*, as bellow:

$$Hip(x) = \frac{e^{x} - e^{-x}}{e^{x} + e^{-x}} \tag{2.8}$$

$$Sig(x) = \frac{1}{1 + e^{-x}} \tag{2.9}$$

Hyperbolic Function takes values between -1 and 1, whereas *Sigmoid Function* takes values between 0 and 1. Since financial markets are nonlinear and have memory suggesting, nonlinear transfer functions are more favored.

There are S neurons and R elements in the input vector displayed on Fig. 2.4. It is a processing unit of the neural network that has one hidden layer.

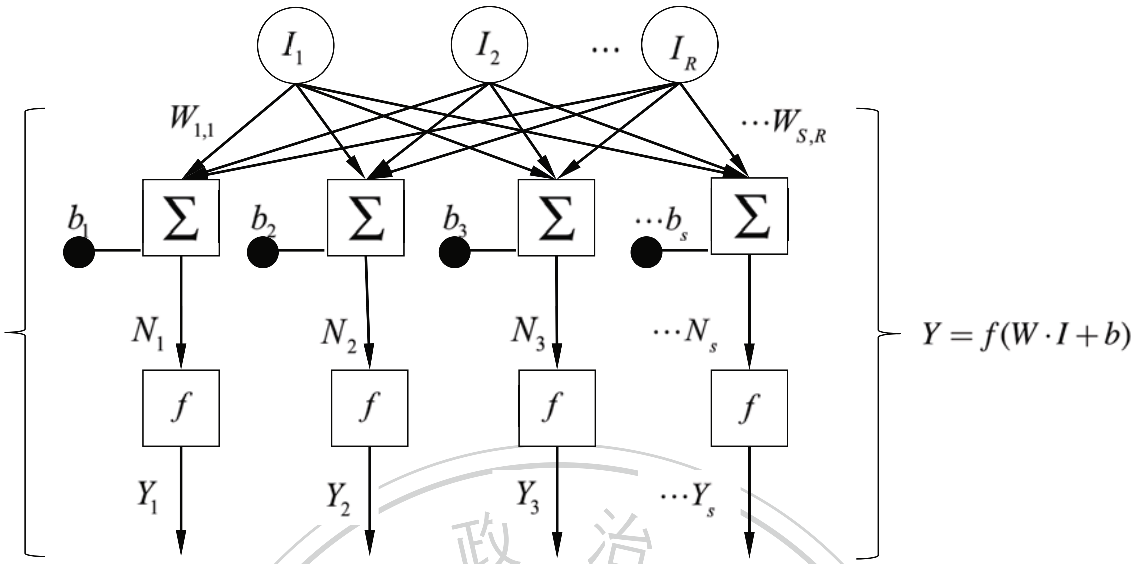

Figure 2.4 A Processing Unit of the Neural Network

A layer of a network is defined in the previous figure. Here, I is the input vector, W is the weight vector, b is the bias vector, N is the summation of the weighted inputs with the bias, f is the transfer function, and Y is the output vector. As equation 2.10 to 2.15 we find

$$I = \begin{bmatrix} I_1 & I_2 & \cdots & I_R \end{bmatrix}^T \tag{2.10}$$

$$W = \begin{bmatrix} W_{1,1} & W_{1,2} & \cdots & W_{1,R} \\ W_{2,1} & \ddots & \cdots & W_{2,R} \\ \vdots & \vdots & \ddots & \vdots \\ W_{S,1} & W_{S,2} & \cdots & W_{S,R} \end{bmatrix} \tag{2.11}$$

$$b = \begin{bmatrix} b_1 & b_2 & \cdots & b_S \end{bmatrix}^T \tag{2.12}$$

$$N = \begin{bmatrix} N_1 & N_2 & \cdots & N_S \end{bmatrix}^T \tag{2.13}$$

$$f = Sig(x), Hip(x) \tag{2.14}$$

$$Y = \begin{bmatrix} Y_1 & Y_2 & \cdots & Y_S \end{bmatrix}^T \tag{2.15}$$

The scalar multiplication of two vectors W and I plus b is defined as N.

$$N = W \bullet I + b = \begin{bmatrix} \sum_{j=1}^{R} W_{1,j} \times P_j + b_1 \\ \sum_{j=1}^{R} W_{2,j} \times P_j + b_2 \\ \vdots \\ \sum_{j=1}^{R} W_{S,j} \times P_j + b_S \end{bmatrix}, j = 1,2,...R \tag{2.16}$$

Finally, N is passed through the transfer function f, which produces the scalar output: Y_i.

$$Y_i = f(N_i) = f(\sum_{j=1}^{R} W_{i,j} I_j + b_i) \tag{2.17}$$

A network can have several layers. In nearly all cases, a three- or four-layer network will do well (Hamid and Iqbal, 2004). We used a four-layer network to simulate the TAIEX Index. The network shown under had R inputs, S neurons in the first hidden layer, K neurons in the second hidden layer, one neuron in the output layer. Transfer function we choose *Hyperbolic Function, Sigmoid Function and Purlin Function*.

Hidden Layer 1 $\qquad f_1 = \dfrac{e^x - e^{-x}}{e^x + e^{-x}}$

Hidden Layer 2 $\qquad f_2 = \dfrac{1}{1 + e^{-x}}$

Output Layer $\qquad f_3 = x$

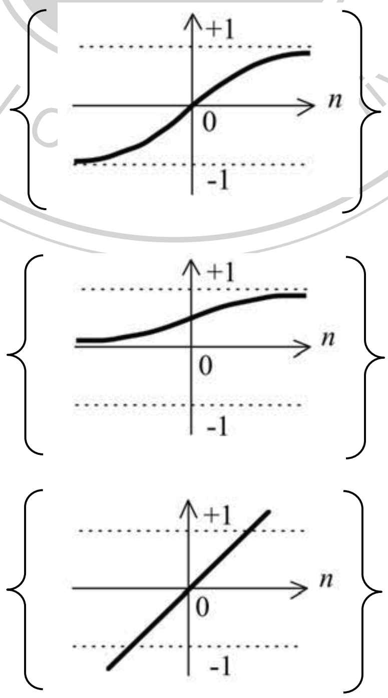

Figure 2.5 Three Transfer functions in this Experimental Design

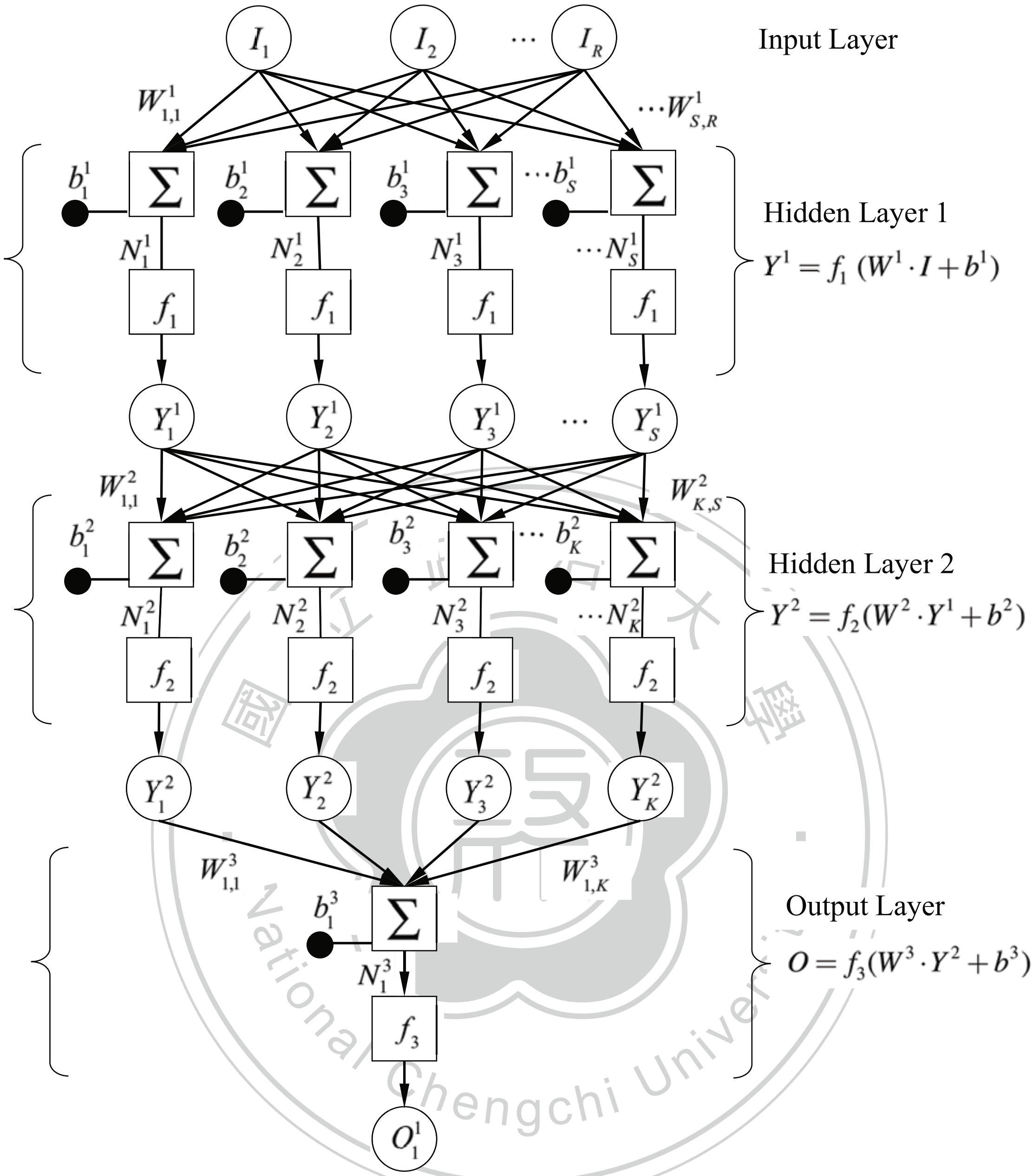

Figure 2.6 A Fully Connected with Four-Layer Feed-forward Network

Note that the outputs of each intermediate layer are the inputs of the following layer. Thus, the hidden layer 2 can be analyzed as a one-layer network with S inputs, K neurons and $S \times K$ weight matrix W^2. This approach can begin with any layer of the network.

2.3 EEMD-Based Neural Network Learning paradigm

A major challenge in neural network learning paradigm is how to make the trained network process good production ability. EEMD provided a good solution. By employing the neural network cross-validation simulate method, neural network can be based on a single series representation for the entire training process, even if the problems are very complex. The result of the neural network learning by single series may be inadequate. The EEMD-Based neural network learning is employed to decompose a time series and approximating it by using decomposed components with a multi variable analysis framework. In this study, we included IMFs and residue components for the final multi-scale learning in the neural model as input variables. Fig. 2.7 presents a Y series that is decomposed into IMFS and then placed into the ANN model as neurons.

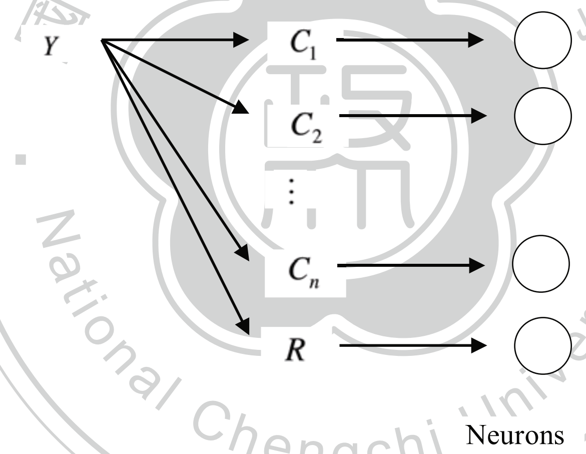

Figure 2.7 EMD Method to Form the Inputs of the Neural Network Model

After decomposition, all the IMFs will show the changing frequencies and amplitudes, which are not the same with any harmonic wave. From the decomposition, the components are analyzed through the ANN method for final multi-scale learning. This simple structure can guarantee a good generalization of the new data. Generally speaking, the EEMD-based neural network learning paradigm consists of the following steps:

1) Adding different ε to the original time series data
2) The original time series are decomposed into various independent IMFs

3) Put all IMFs as the input layer into an artificial neural network.
4) Process matrices by mapping original series minimum and maximum values to [-1 1].
5) Dividing the database into two sets: the training set and the testing and validation set.
6) Creating an ANN structure with two hidden layers with different neurons. Here we set 3 numbers of neurons to compare the forecasting results.
7) The number of iteration has infinite epochs.
8) Reverse the processing of result to get predicts.

Since hidden neurons may perform better between half the number of input variables and twice that number. We chose neurons between those favorable intervals. Table 2.1 consists of the experimental parameter's variable sets. We gave all the components a different code.

Table 2.1 Experimental Code Set

ε	**Hidden layer 1**	**Hidden layer 2**	**Output layer**	**Code**
	8	5	1	A
0.1	8	10	1	B
	8	16	1	C
	8	5	1	D
0.2	8	10	1	E
	8	16	1	F
	8	5	1	G
0.3	8	10	1	H
	8	16	1	I

3. Index Options

An option is a stipulated privilege of buying or selling a stated security, commodity or property at a given price (strike price) within a specified time. The important point to emphasize is that an option gives the hold the “right” to do something but doesn't have to exercise this right. This distinguishes option from forward and futures, where the holder to buy or sell the underlying security. If the underlying security is an index we called index option. When exercised, settlement is made by cash payment, since physical delivery is not possible.

Types of options

There are two types of options, which are calls and puts: A call gives the holder the right to buy an asset at a certain price within a specific period of time. Calls are similar to having a long position on a stock. Buyers of calls hope that the stock will increase substantially before the option expires; a put gives the holder the right to sell an asset at a certain price within a specific period of time. Puts are very similar to having a short position on a stock. Buyers of puts hope that the price of the stock will fall before the option expires.

Table 3.1 Four types of participants in options markets

Roles	Type	Component	Risk	Reward	Market Trend
Buyers	Call	Long call	Limited to the premium paid up front for the option	Unlimited as the market rallies	Bullish
Sellers	Call	Short call	Unlimited as the market rises	Limited to the premium received for selling the option	Bearish
Buyers	Put	Long put	Limited to the net premium paid for the option	Unlimited as the market sells off	Bearish
Sellers	Put	Short put	Unlimited in a falling market	Limited to the premium received for selling the put option	Bullish

People who buy options are called holders and those who sell options are called writers; furthermore, buyers are said to have long positions, and sellers are said to have short positions. When the market trends are increasing, buying a call is a position on an underlying stock price, thus making a bullish decision. The investor has the opportunity to participate in the rise of the stock's value for the term of that contract with a predetermined risk. Most investors will look to sell their contract at a profit, while others may intend to exercise their right and purchase the underlying shares. The main benefit of buying a call is the limited risk of capital. The investor has a much smaller cash layout, with a limited downside loss, and unlimited upside gain. On the flip side, the option investor does not have the same rights of the individual shareholder such as dividends and voting rights. On the other hand, when the market trend is diminishing, buying a Put is a bearish decision to be placed in a position on the underlying stock value.

The Premium

Buyers need to pay premium to get option positions. To determine the premium of option is called pricing option. The formulas derived by Black and Scholes (1973) are the most important model for pricing options. The model is based on some assumption such as geometric Brownian motion of stock prices movement, continuous trading with no dividends and taxes applied to the stocks, and that the market is frictionless. Based on random walk hypotheses of stock movement, the volatility of stocks is estimated based on historical data. The pair of formulas for call and put, respectively, in Equation 3.1 constitutes the Black Scholes model.

$$c = SN(d_1) - Xe^{-rT}N(d_2);\ \ p = Xe^{-rT}N(-d_2) - SN(-d_1) \tag{3.1}$$

Where $d_1 = \dfrac{\ln(S/X) + (r + \frac{1}{2}\sigma^2)T}{\sigma\sqrt{T}}$, $d_2 = d_1 - \sigma\sqrt{T}$

c the call price, p the put price;

$N(\ )$ the cumulative probability function

S the current underlying asset price;

X the strike price;

T the time to maturity (in year);

σ the volatilities of underlying asset;

r the short term risk free interest rate.

According to Black Scholes model, there are five major factors influencing index option premium. The first factor is changing the price of the Taiwan stock exchange capitalization weighted stock index (TAIEX). These price changes have opposite effects on calls and puts. It can increase or decrease the value of an option. The second factor is the strike price because it determines whether or not an option has any intrinsic value. Premium generally increases as the option becomes further in the money[1], and decreases as the option becomes deeply out of the money[2]. Only in the money options have intrinsic value, representing the difference between the current price of the underlying security and the option's strike price. Another factor that effects the index is the period of time until expiration, that is, as expiration approaches, both put and call option will decrease. This effect is most noticeable with the money[3] option. Not only will time influence the index but also the volatility of the TAIEX will do so. Volatility is a statistical measure of how much variation a stock exhibits as it fluctuates around its mean. Higher volatility estimates reflect greater expected fluctuations in underlying price levels, however, more volatility means a greater chance that the option will move. This expectation generally results in higher option premiums for both puts and calls. The last factor is current risk-free interest rate. This impact reflects the shares related to the cost of carry in an underlying security. Cost of carry means costs incurred as a result of an investment position.

We buy and sell options within the same trading day because all positions are usually closed before the market close for the trading day. Since this experiment is done in one day trading,, short-term risk free interest rate is not significant, so it can be ignored. We only focus on the underlying asset price and volatility. They will appear in our predicted results. In this research, we collected data from September 2010, 21 trading days, TAIEX options tick-by-tick data from Taiwan Future Exchange (TAIFEX) to test our trade strategy model. TAIEX options were

---

[1] In the money: For a call option, the strike price is below the market price and put option is opposite if strike price is above the market price.

[2] Out of the money: For a call option, the strike price is higher than the market price and put option is opposite if strike price is below the market price.

[3] At the money: If the strike price of the option equals the market price of the underlying security.

introduced on 24th December 2001. Table 3.2 and 3.3are the trade volume statistics and specification for the TAIEX options. We can see TAIEX option trading is very alive and kicking in Taiwan

Table 3.2 TAIEX Option Annual Trading Volume (Thousand)

Year	**2001**	**2002**	**2003**	**2004**	**2005**	**2006**	**2007**	**2008**	**2009**	**2010**
Volume	5	1,566	21,720	43,825	80,097	96,930	92,586	92,757	72,083	95,667

Table 3.3 The Specification for TAIEX options

Item	Description
Underlying Index	Taiwan Stock Exchange Capitalization Weighted Stock Index (TAIEX)
Ticker Symbol	TXO
Exercise Style	European
Multiplier	NTD 50 (per index point)
Expiration Months	Spot month, the next two calendar months, and the next two quarterly months
Strike Price Interval	1. Strike price below 3,000 points: 50 index points in the spot month and the next two calendar months, 100 index points in the two quarterly months 2. Strike price at or above 3,000 but below 10,000 points: 100 index points in the spot month and the next two calendar months, 200 index points in the two quarterly months 3. Strike price at or above 10,000 points: 200 index points in the spot month and the next two calendar months, 400 index points in the two quarterly months
Strike (Exercise) Price	For listing series of new expiration months or series with new strike prices for existing expiration dates, the TAIFEX shall, based on the previous business day's closing price of the underlying index, consecutively introduce series with new strike prices according to the strike price interval above, till the following conditions are satisfied: 1. For the spot month and the next two calendar months, the highest and lowest strike prices should cover 15 percent above and below the underlying index; 2. For the next two quarterly months, the highest and lowest strike prices should cover 20 percent above and below the underlying index.
Daily Price Limit	+/- 7% of previous day's closing price of the underlying index
Position Limit	1. Any investor's aggregate open same-side positions in the Contract for various delivery months at any time shall not exceed the limit standards announced by the TAIFEX 2. Institutional investors may apply for an exemption from the above limit on trading accounts for hedging purpose 3. These position limits are not applicable to omnibus accounts
Trading Hours	1. 08:45AM - 1:45 PM Taiwan time Monday through Friday of the regular Taiwan Stock Exchange business days. 2. 08:45AM-1:30 PM on the last trading day for the delivery month contract.
Last Trading Day	The third Wednesday of the delivery month
Expiration Date	The same day as the last trading day
Final Settlement Price	The average price of the underlying index disclosed within the last 30 minutes prior to the close of trading on the final settlement day.
Settlement	Cash settlement. An option that is in-the-money and has not been liquidated or exercised on the expiration day shall, in the absence of contrary instructions delivered to the Exchange by the Clearing Member representing the option buyer, be exercised automatically

Source: Taiwan Future Exchange

Our expiration month is spot month. The reason the expiration month was used is that investors prefer to trade spot month for trade volume, and we design a day trading strategy. Figure 3.1 is a real stock price in September, and the color gray is displays the strike price we chose.

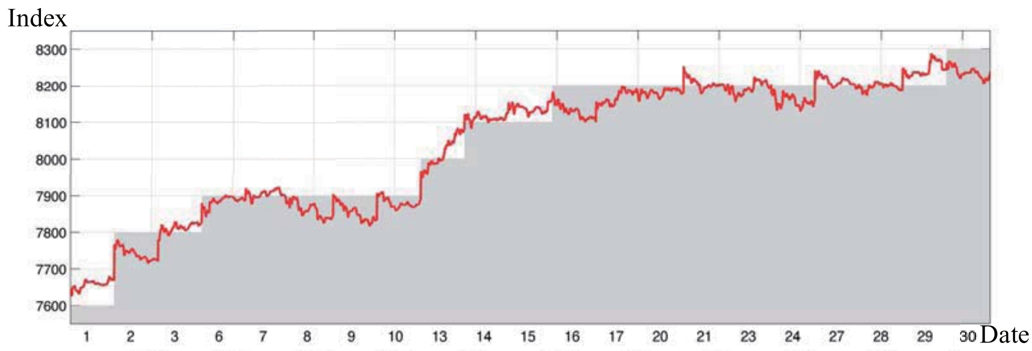

Figure 3.1 The Strike Price and The Real Index in September

We rounded the open spot price to the nearest hundred then use it as the strike price we will choose for trading.

Fig. 3.2 is a historical call and put option according to the price we chose.

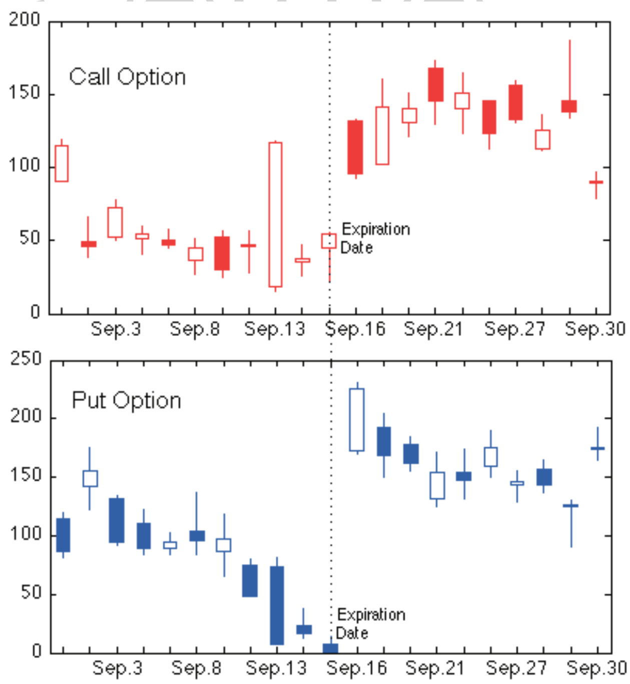

Figure3.2 Options Price on Spot Month

If the close price was lower than the open price, the candlestick will be filled with black, otherwise the candlestick is filled with white, which means the close price was higher than open price. Note that the expiration date is on Sep. 15th, the put option price is close to zero. Because the time value of an option tends to increase as time gets closer to its expiration date because a longer-lived option has a greater chance of going deeper in the money before it expires. When expiration time value is zero and is the only value, the option has is its intrinsic value.

4. Algorithmic Trading

Algorithmic Trading (AT) was originally developed for use by the buy-side to manage orders and to reduce market impact by optimizing trade. Hence, AT may be defined as an electronic trading whose parameters by strict adherence to a predetermined set of rules. Investors and traders use AT to chase specific execution outcome and determine the timing, underlying price, quantity even others routing orders.

In this research, we decomposed the Taiwan Weighted Stock Index for the period of Aug. 30, 2010 for half a day and 135 minutes as the input of our test, and time delay for one minute as the target. The ANN structure is used to forecast the performance of next trading day.

4.1 Moving FK Indicator

Zhang (1993) provided an indicator to determine the timing for trade and is known as Forward-Calculator % K (FK). Indicators are in the range from 0 to 1. As used in equation 4.1

$$FK_n = \frac{C_n - C_{\min}}{C_{\max} - C_{\min}} \tag{4.1}$$

C_n is an observation data on trading period; C_{max} has the largest observation data and C_{min} is the lowest one. The following data are from Sep. 02, 2010 Taiwan stock exchange index for every minute. According to our definition, trade signals are shown over 0.8 and under 0.2.

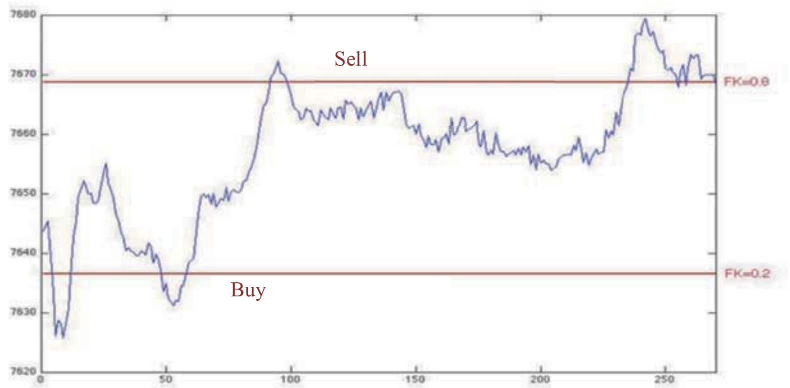

Figure 4.1 Sep. 02,2010 TWSE Index and FK Indicator (min)

In this research, we improved a moving FK indicator to define any time period.

$$FK_n^k = \frac{P_n - C_{\min}^k}{C_{\max}^k - C_{\min}^k} \tag{4.2}$$

P_n is the predicted data on trading period, and $C_{\max}^k$ is the largest one, and $C_{\min}^k$ is the lowest one on k^{th} period.

Considering the trading period, the value of FK indicator will move toward 0 when the stock price is rising. Falling stock price will run in the opposite direction that is, it will move toward the value of 1.Through the FK indicator, investors will be able to grasp the position and trend of the General Index. In other words, the value of FK indicator predicts the corresponding level of the current stock index at a specific point of time in the future. Investors, therefore, can set up a threshold value to determine the timing for trade. A buy-trade can be executed when the value of FK indicator falls below a specific threshold and a sale-trade can be considered when the value has exceeded the established threshold. In this research, the system will output a "buy" signal (value = 1), but when the value of FK indicator is under the lower threshold, the "sell" signal (value = -1) will appear. In other situations, the system will output a "hold" (value = 0) signal.

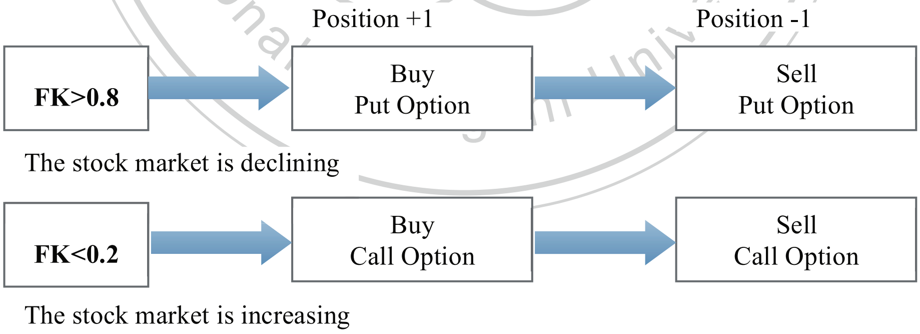

Figure4.2 FK signal and reaction of strategies

Fig. 4.2 is a kind of reaction when FK signal reaches a threshold. When the market is open, our position limit is 5.

4.2 Process

For all predictions are from the following steps:

1. Read the previous day stock price
2. Decomposed stock price to differ IMFs.
3. Simulate IMFs by using ANN method to predict next minute price.
4. Transform predict result to FK indicator.
5. Determine FK value if it shows buy or sell signal.
6. Confirm trading time and make buy or sell Call-Put decisions.
7. After one day trading, count profit result.

The main purpose of this study is to provide an alternative and promising scheme for the ANN prediction. We use moving window concept in our model to predict the leading minute data. We employ the method of rolling window which its length is fixed to 135 minutes[4]. After simulating one window, we move it to next minute and replace the prediction data with the real one, and repeat the process for the whole month.

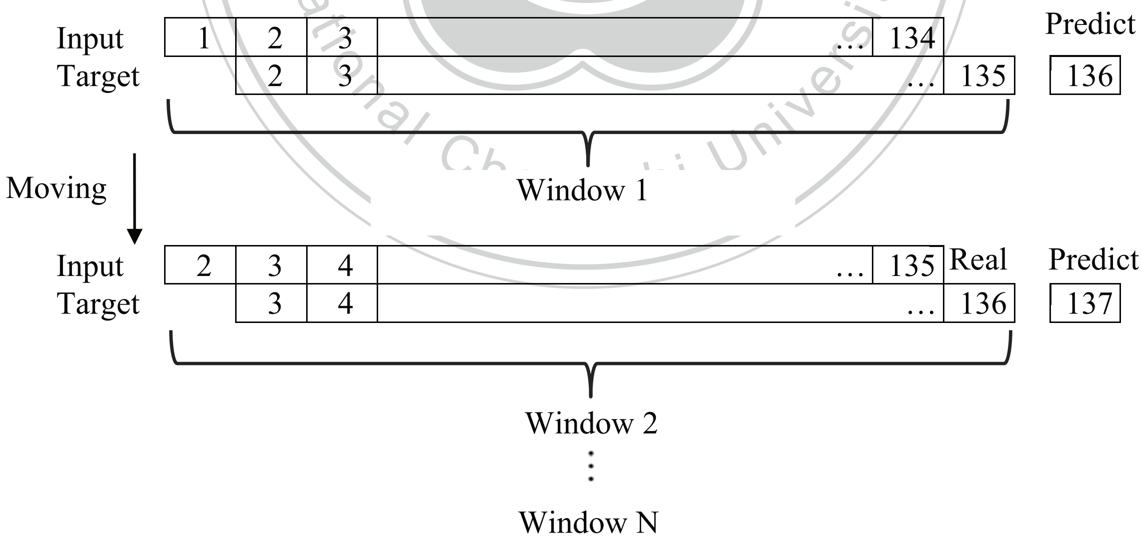

Figure 4.3 Training Network by Moving Window Process

[4] Appendix A shows why we set 135 minutes in our moving window

There are three trade combinations from our trade strategy: buy-sell-put, buy-sell-call and mix-put-call, seen below:

1	Buy Put	**-1**	Buy Call	**1**	Buy Put
-1	Sell Put	**1**	Sell Call	**-1**	Buy Call
1	Buy Put	**-1**	Buy Call	**-1**	Sell Put
-1	Sell Put	**1**	Sell Call	**1**	Sell Call

Figure 4.4 Combination 3 Trades from Buy/Sell Options

If the market is in a bullish condition, we can always get a -1 signal, then our program would show "buy put" instruction. However, on the other side, if the market is in a bearish condition, our program would show "buy call" instruction. If the stock price does not have evident direction, then our program would let us buy the "mix-put-call" strategy. The strategy also called "Straddle" , which is a neutral strategy in option trading that involves the simultaneous buying of a put and a call of the same strike price. Beside the moving FK indicator means, we could know whether if the in-time price is high or low in the market.

FK indicator present

The whole month's FK indicator results is as follows. Red means the stock price is very high in that time, and blue means price is low at that time. As a result, it draws a filled contour plot of FK result matrix, whereas the matrix is interpreted as "heights" with respect to the *x-y* plane. The market price increased in Sep. 8 and Sep.13, so the moving FK values are over 0.5 on those day. Though the whole picture seems likely, but we still can see different colors in detail, and that may affect our trading structure.

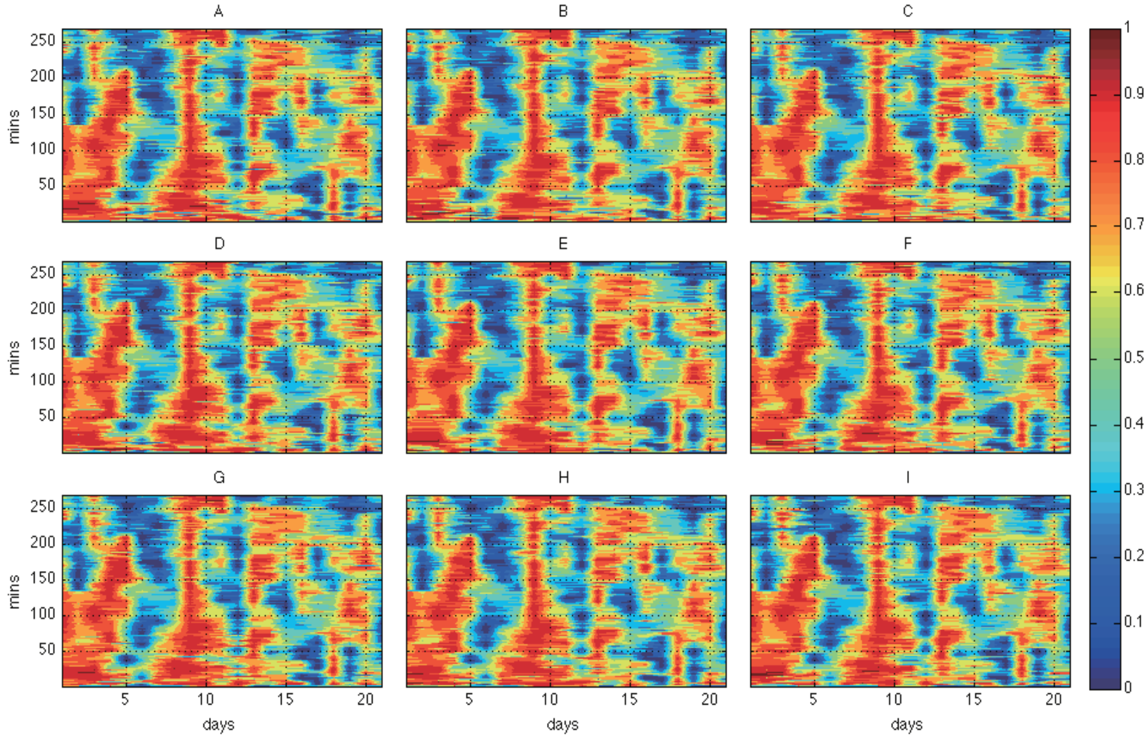

Figure 4.5 FK Results in September 2010 with Different Experimental Set

4.3 Performance

Predict result

We want to measure forecast accuracy and remove the sign of prediction errors, so we calculate the root mean square errors (RMSE) where

$$RMSE = \left[\frac{1}{N} \sum_{t=1}^{N} (\hat{x}_t - x_t)^2 \right]^{\frac{1}{2}} \tag{4.3}$$

And the mean of absolute errors (MAE) where

$$MAE = \frac{1}{n} \sum_{t=1}^{n} \left| \hat{x}_t - x_t \right| \tag{4.4}$$

These two equations help determine the index forecast compared with reality TWSE index. Table 4.1 categorizes RMSE's and MAE's Predict result during September 2010. MAE is counted per day; RMSE is counted whole horizons, and is seen as follows:

Table 4.1 RMSE and MAE of Predict result on September 2010

Condition	**ε=0.1**			**ε=0.2**			**ε=0.3**		
	8-5-1	**8-10-1**	**8-16-1**	**8-5-1**	**8-10-1**	**8-16-1**	**8-5-1**	**8-10-1**	**8-16-1**
Code	**A**	**B**	**C**	**D**	**E**	**F**	**G**	**H**	**I**
RMSE	**7.205**	**6.639**	**6.988**	**6.776**	**6.316**	**6.945**	**6.643**	**6.296**	**6.842**
Rank RMSE	**9**	**3**	**8**	**5**	**2**	**7**	**4**	**1**	**6**
Date		**MAE**			**MAE**			**MAE**	
0901	3.057	3.143	3.199	3.186	2.963	3.451	2.910	3.186	3.364
0902	4.135	4.597	4.825	5.027	4.841	4.482	3.908	4.067	4.146
0903	3.983	3.819	4.480	3.907	3.835	3.905	4.087	3.754	4.092
0906	3.244	3.125	3.180	2.899	2.718	3.514	3.680	3.233	3.269
0907	2.861	2.494	2.873	2.768	2.350	2.669	2.419	2.818	2.616
0908	3.939	3.495	3.565	3.766	3.654	3.640	3.696	3.328	3.472
0909	3.973	3.603	3.742	3.987	3.381	3.390	3.980	3.370	4.107
0910	4.321	3.327	3.919	3.879	3.639	3.924	3.425	3.839	3.762
0913	5.335	5.410	5.848	4.835	5.045	5.778	4.915	5.345	5.197
0914	3.701	3.628	3.429	3.504	3.007	3.705	3.492	3.180	3.802
0915	3.884	3.540	4.039	3.193	3.469	3.728	3.101	3.303	3.760
0916	3.318	3.373	3.324	3.184	3.278	3.620	4.189	3.626	3.783
0917	4.837	4.663	4.673	4.486	3.856	4.254	3.875	4.327	3.770
0920	2.597	2.655	2.841	2.776	2.680	2.564	2.793	2.815	2.463
0921	3.550	2.746	3.404	3.219	3.131	3.554	3.686	3.289	3.682
0923	3.213	2.974	3.517	3.335	3.343	3.455	3.540	3.444	3.318
0924	4.359	4.409	4.551	4.353	4.138	4.611	4.512	4.066	4.453
0927	3.648	3.356	3.954	3.630	3.159	3.778	3.343	3.177	3.982
0928	2.875	2.491	2.549	2.865	2.621	2.707	2.637	2.328	2.622
0929	4.545	4.591	4.641	4.637	4.410	5.141	4.741	4.398	5.360
0930	3.647	3.132	3.365	3.361	3.376	3.331	3.480	3.366	3.405
Mean of MAE	**3.763**	**3.551**	**3.806**	**3.657**	**3.471**	**3.771**	**3.639**	**3.536**	**3.735**
Rank MAE	**7**	**3**	**9**	**5**	**1**	**8**	**4**	**2**	**6**

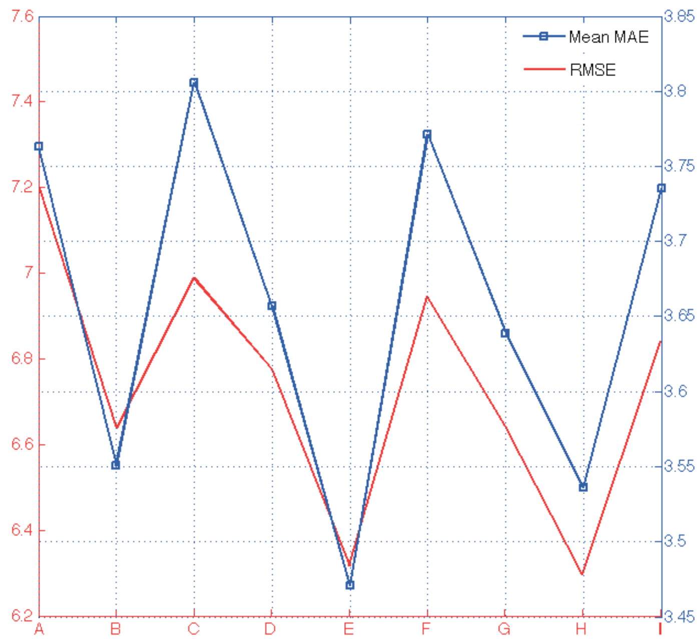

Figure 4.6 Results of RMSE and mean MAE

Fig. 4.6 displays RMSE and MAE results of different groups. We find that group B, E and H have better performance in the predicted results. It seems that the best neuron numbers in our case is 8-10-1.

The Daily profit

In this study, we discussed four cases of different threshold to see the result.

Table 4.2 The threshold of FK indicator

	Bellow	Above
Case 1	0.15	0.85
Case 2	0.20	0.80
Case 3	0.25	0.75
Case 4	0.30	0.70

The final monthly return %

The monthly return performance shown in fig 4.7 allows us to be found if we set strict threshold numbers, like 0.15 and 0.85, then the trade signal will present the real market near extreme prices. The return is also in direct proportion to the correct prediction set. Set B, E and H in threshold at 0.15 and 0.85 have the largest return at 6.0 %, 6.2 % and 6.1 %.

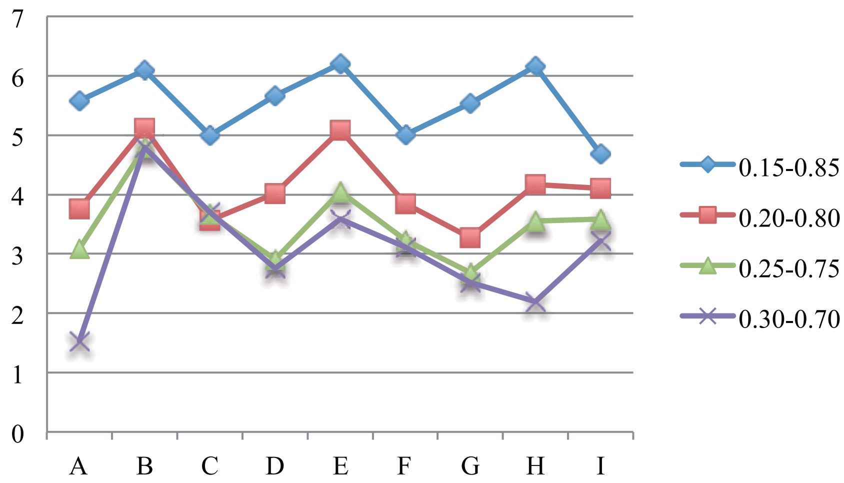

Figure 4.7 Monthly return % and FK threshold

Beside our day trading strategy, there are obviously many different strategies to trading options. The option term is the period during which an option is active. During this term, the option changes in value based on fluctuations in the price of the stock it represents. If you make good predictions on the stock market, option is the one product that you can gain the most profit.

Appendix C provides a method to discuss the cost of trade because the transaction fee varies by different brokers. Further more, we can add the transaction fee and tax to calculate the real profit.

5. Conclusion

As researchers and investors strive to outperform the market, today we can forecast the stock market prices by the uses of neural networks. Neural networks appear to be the best modeling method currently available as they capture nonlinearities in the system without human intervention. Continued work on improving neural network performance may lead to more insights in the chaotic nature of the systems they model. In this study, we used EEMD-base ANN learning algorithm to forecast stock price per minute. It displays a good prediction on stock price and also adapt to option market. In our study, we provide a method for option day trading. Besides observing the monthly performance, researchers can also use this study to show how the good prediction can earn much profit. Day trading options are one way to take advantage of minimal moves in an underlying stock price. Based on the nature of options, they can provide a much greater return than day trading stocks. By adding options to your day trading portfolio, it also comes with increased risks compared to buying and selling stocks.

APPENDIX A.

We tested the different length of our moving window to check what minutes would be the best adjusted, so we decided to choose 30, 60, 90 and 135 minutes for training the neural network.

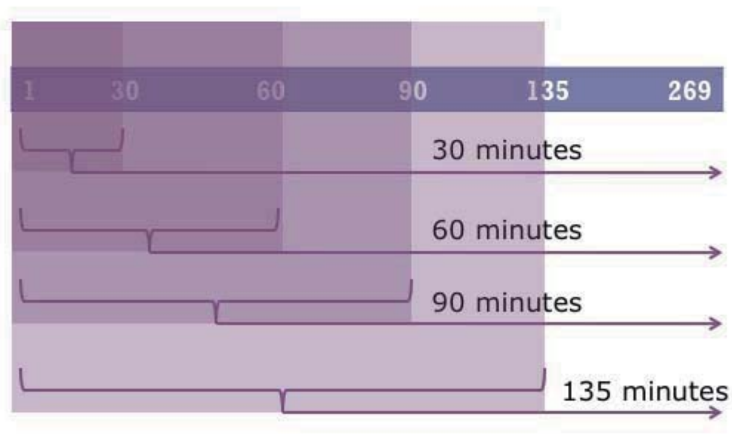

Figure A.1 Moving Window Test

As the predict result, we compared the predict data and real data, and count the variance of errors. We found when the time window is 135 minutes has the best performance.

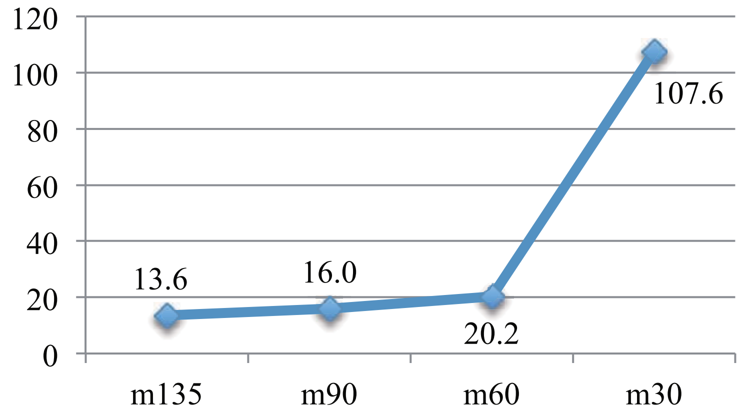

Figure A.2 The Variance of Error

APPENDIX B.

Table A.1 Profits of FK threshold=0.15-0.85, trading position=5

Position=5	ε=0.1			ε=0.2			ε=0.3		
	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1
Monthly Return %	**5.580**	**6.099**	**4.999**	**5.662**	**6.203**	**5.011**	**5.537**	**6.165**	**4.684**
Monthly Expense	**8665.5**	**8445.1**	**8845.6**	**8468**	**7746.4**	**8767.5**	**8695.9**	**8942.6**	**9095.5**
SUM	483.5	515.1	442.2	479.5	480.5	439.3	481.5	551.3	426
1-Sep	-54	-62	-37	-46	-35	-36	-40	-61	-51
2-Sep	87	79	91	94	109	106	88	116	88
3-Sep	-15	0	-26	0	0	0	-20	0	-2
6-Sep	-32	-32	-51	-13	-50	-50	-34	-68	-69
7-Sep	36	34	33	32	36	32	38	30	33
8-Sep	-6.5	0	1	-1.5	0	0	-8	0	-1
9-Sep	100	96	95	93	99	94	98	109	102
10-Sep	78	35	39	28	41	36	9	62	47
13-Sep	11	0	0	5	0	0	0	0	0
14-Sep	-19.5	-29.5	-18.5	-37.5	-10	-46	-10	-11	-31.5
15-Sep	-17.5	-5.4	-4.3	-15.5	-9.5	-12.7	15.5	-14.7	-9.5
16-Sep	59	60	59	60	66	63	50	64	68
17-Sep	-30	-25	-52	-74	-80	-53	-80	-70	-51
20-Sep	34	55	46	49	41	68	59	58	44
21-Sep	129	172	148	158	142	118	173	164	142
23-Sep	24	41	38	35	31	33	41	70	40
24-Sep	0	7	9	7	6	12	3	-4	9
27-Sep	47	54	51	59	47	52	57	59	50
28-Sep	64	43	42	47	58	65	56	50	44
29-Sep	-13	-10	-11	0	-11	-39	-14	-4	-21
30-Sep	2	3	-10	0	0	-3	0	2	-5

Table A.2 Profits of FK threshold=0.2-0.8, trading position=5

Position=5	ε=0.1			ε=0.2			ε=0.3		
	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1
Monthly Return %	**3.760**	**5.121**	**3.562**	**4.021**	**5.088**	**3.847**	**3.271**	**4.173**	**4.099**
Monthly Expense	**9804**	**9225**	**10102.5**	**9665.4**	**8957.9**	**9602.6**	**9696.3**	**9774.6**	**9851.7**
SUM	368.6	472.4	359.9	388.6	455.8	369.4	317.2	407.9	403.8
1-Sep	-57	-55	-59	-57	-64	-41	-68	-63	-63
2-Sep	94	83	91	92	109	94	88	93	88
3-Sep	-15	0	-26	0	0	-17	-18	0	0
6-Sep	-67	-62	-88	-87	-84	-85	-49	-91	-96
7-Sep	34	34	28	34	36	24	38	30	36
8-Sep	-5.5	-5.5	1	-0.5	0	0	-5.5	0	-1
9-Sep	96	98	97	97	96	94	78	95	96
10-Sep	77	79	80	73	77	58	53	76	78
13-Sep	11	0	0	5	0	0	-8	0	0
14-Sep	-50	-51.5	-36.5	-47	-28.5	-42.5	-40.5	-50.5	-54.5
15-Sep	-12.9	-9.6	-6.6	-15.9	-8.7	-6.1	-3.8	-12.6	-5.7
16-Sep	54	58	59	59	63	60	49	65	55
17-Sep	-48	-16	-61	-59	-52	-53	-74	-58	-50
20-Sep	47	56	45	59	55	64	56	60	53
21-Sep	100	172	131	134	142	117	115	162	160
23-Sep	21	28	37	36	32	44	39	31	41
24-Sep	12	8	20	-1	6	13	-13	4	7
27-Sep	56	50	44	56	54	51	57	54	48
28-Sep	26	21	23	20	45	34	23	28	26
29-Sep	-6	-21	-7	-5	-22	-37	-1	-14	-25
30-Sep	2	6	-12	-4	0	-2	2	-1	11

Table A.3 Profits of FK threshold=0.25-0.75, trading position=5

Position=5	ε=0.1			ε=0.2			ε=0.3		
	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1
Monthly Return %	**3.086**	**4.791**	**3.670**	**2.905**	**4.054**	**3.230**	**2.685**	**3.556**	**3.588**
Monthly Expense	**10120**	**9701.5**	**10447**	**10160.9**	**9754.9**	**10035.1**	**10415.5**	**10026**	**10265.4**
SUM	312.3	464.8	383.4	295.2	395.5	324.1	279.7	356.5	368.3
1-Sep	-62	-60	-67	-67	-65	-35	-67	-63	-66
2-Sep	84	83	96	81	78	94	84	83	85
3-Sep	-15	-19	-26	0	3	-17	-20	-25	0
6-Sep	-87	-69	-75	-89	-85	-87	-50	-92	-91
7-Sep	34	30	29	25.5	27	24	31.5	24	36
8-Sep	-16.5	-5.5	-5	-11	0	3	-13.5	0	-1
9-Sep	77	96	95	74	96	96	80	96	94
10-Sep	78	80	80	73	77	63	58	75	77
13-Sep	11	0	3	5	0	0	-10	0	0
14-Sep	-49.5	-37.5	-45	-37	-48.5	-42	-40	-40.5	-50.5
15-Sep	-19.7	-7.2	-2.6	-12.3	-10	-16.9	-9.3	-14	-7.2
16-Sep	57	49	60	59	66	51	42	64	46
17-Sep	-13	0	-65	-64	-52	-49	-80	-56	-52
20-Sep	35	44	48	46	52	41	43	40	51
21-Sep	79	173	131	106	155	75	107	163	128
23-Sep	21	29	24	36	34	43	44	24	49
24-Sep	14	18	17	6	4	14	1	8	13
27-Sep	55	54	49	56	56	50	56	57	46
28-Sep	21	22	21	18	30	24	22	24	22
29-Sep	10	-21	13	-5	-21	-19	-8	-15	-24
30-Sep	-1	6	3	-5	-1	12	9	4	13

Table A.4 Profits of FK threshold=0.3-0.7, trading position=5

Position=5	ε=0.1			ε=0.2			ε=0.3		
	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1	8-5-1	8-10-1	8-16-1
Monthly Return %	**1.532**	**4.792**	**3.709**	**2.757**	**3.590**	**3.110**	**2.520**	**2.197**	**3.224**
Monthly Expense	**10627.5**	**10173**	**10638**	**10682.5**	**10314**	**10446**	**10543**	**10620**	**10466.5**
SUM	162.8	487.5	394.6	294.5	370.3	324.9	265.7	233.3	337.4
1-Sep	-61	-50	-39	-62	-60	-30	-67	-62	-61
2-Sep	62	83	63	82	64	94	85	64	82
3-Sep	-40	-19	-49	-20	3	-34	-43	-85	0
6-Sep	-91	-65	-69	-88	-93	-86	-49	-88	-66
7-Sep	31	18	28	22.5	27	28	28	24	36
8-Sep	-22	-13	-12	-9	-8	-8	-5.5	-7.5	-14
9-Sep	53	93	89	74	93	84	73	95	77
10-Sep	46	90	66	70	79	45	58	88	77
13-Sep	11	0	3	5	0	0	-10	0	0
14-Sep	-44	-37.5	-31.5	-28.5	-39.5	-31.5	-34.5	-38	-42
15-Sep	-19.2	-5	0.1	-11.5	-7.2	-13.6	-9.3	-10.2	-6.6
16-Sep	40	39	59	56	61	38	34	50	39
17-Sep	-23	-6	-47	-57	-46	-32	-69	-62	-59
20-Sep	29	41	46	42	41	50	49	47	48
21-Sep	30	152	128	79	102	63	92	91	80
23-Sep	20	20	42	33	35	42	31	34	49
24-Sep	21	19	15	16	22	17	11	13	14
27-Sep	56	55	56	60	54	49	55	57	47
28-Sep	17	23	14	20	27	29	19	21	23
29-Sep	33	37	27	-1	-8	7	5	-9	3
30-Sep	14	13	6	12	24	14	13	11	11

APPENDIX C.

The trade Cost

Since we didn't have "bid and ask" premium price, our data included many trade ticks in one minute. As the tick data were recorded per one thousandth per minute, we still couldn't figure out the best trade premium cost. We used the Counter Daily Potential (CDP) average price to modify the minutes trading premium. The CPD average price is determined from the previous stock price.

$$CDP = \frac{High + Low + 2Close}{4} \tag{4.5}$$

In our case, the "*high*" price is most expensive price and *"low"* price is the cheapest price in a certain minute. *Close* price is counted twice because the close price was thought as more weighted in a trading volume. Furthermore, the transaction fee and tax were counted as NT$100 per position trading. If we traded 5 positions a day, the cost of transaction fee and tax would be NT$1,000.

Reference

Abu-mostafa Y. S. and Atiya A. F., Introduction to Financial Forecasting, Applied Intelligence, 1996, 6: 205-213.

Black F. and Scholes M.,The Pricing of Options and Corporate Liabilities, The Journal of Political Economy,1973, vol. 81,No. 3:637-654.

Chan M. C., Wong C. C., Lam C. C., Financial time series forecasting by neural network using conjugate gradient learning algorithm and multiple linear regression Weight Initialization, 2000, Citeseer.

Djennas Me., Benbouziane M. and Djennas Mu., An Approach of Combining Empirical Mode Decomposition and NeuralNetwork Learning for Currency Crisis Forecasting, Politics and Economic Development, ERF 17th annual conference, 2011, Turkey.

Huang N.E., Shen Z., and Long S. R., The empirical mode decomposition and the hilbert spectrum for onlinear and non-stationary time series analysis, Process of the Royal Society of London, 1998, A454: 903–995.

Huang N. E., Wu M. L. Qu W. D., Long S. R., Shen S. P.and Zhang J. E., Applications of Hilbert–Huang transform to non-stationary financial time series analysis, Appl. Stochastic Models Bus. Ind., 2003: 245:268

Hamid S. A. and Iqbal Z., Using neural networks for forecasting volatility of S&P 500 Index futures prices, Journal of Business Research, 2004, 57: 1116-1125

Kaastra I., Boyd M., Designing a neural network for forecasting financial and economic time series, Neurocomputing, 1996, 10: 215-236

Klevecka, I., Lelis J., Pre-Processing of Input Data of Neural Networks: The Case of Forecasting Telecommunication Network Traffic, Telenor ASA, 2008

Lin, T. W. and Yu, C. C., Forecasting stock market with neural networks, SSRN Working Paper, 2009

Mendelsohn L., Preprocessing data for Neural Networks, Tech Anal Stocks Commod, 1993:52-58

Wu, Z. and Huang, N.E., Ensemble empirical mode decomposition: a noise-assisted data analysis method, Centre for Ocean Land Atmosphere Studies. Technical Report, 2004, 193: 51

Yu. L., Wang, S. Y. and Lai, K. K., Forecasting crude oil price with an EMD-based neural network ensemble learning paradigm, Energy Economics,2008, 30: 2623-2635

Yu L., Wang S. Y., Lai K. K., Wen F. H., A multiscale neural network learning paradigm for financial crisis forecasting, Neuro computing, 2010, 73:716-725

Zhang, X., Lai, K.K., and Wang, S. Y., A new approach for crude oil price analysis based on Empirical Mode Decomposition, Energy Economics, 2008, 30: 905-918