Expert Systems with Applications 37 (2010) 4742–4748

Contents lists available at ScienceDirect

Expert Systems with Applications

journal homepage: www.elsevier.com/locate/eswa

Adapted Neuro-Fuzzy Inference System on indirect approach TSK fuzzy rule base for stock market analysis

Akbar Esfahanipour, Werya Aghamiri *

Department of Industrial Engineering, Amirkabir University of Technology (Polytechnic of Tehran), Tehran, Iran

ARTICLE INFO

Keywords:
FCM clustering
Fuzzy Rule Based Systems
Forecasting
Stock market
Forecasting accuracy
ANFIS

ABSTRACT

Nowadays because of the complicated nature of making decision in stock market and making real-time strategy for buying and selling stock via portfolio selection and maintenance, many research papers has involved stock price prediction issue. Low accuracy resulted by models may increase trade cost such as commission cost in more sequenced buy and sell signals because of insignificant alarms and otherwise bad diagnosis in price trend do not satisfy trader's expectation and may involved him/her in irrecoverable cost. Therefore, in this paper, Neuro-Fuzzy Inference System adopted on a Takagi–Sugeno–Kang (TSK) type Fuzzy Rule Based System is developed for stock price prediction. The TSK fuzzy model applies the technical index as the input variables and the consequent part is a linear combination of the input variables. Fuzzy C-Mean clustering implemented for identifying number of rules. Initial membership function of the premise part approximately defined as Gaussian function. TSK parameters tuned by Adaptive Nero-Fuzzy Inference System (ANFIS). Proposed model is tested on the Tehran Stock Exchange Indexes (TEPIX). This index with high accuracy near by 97.8% has successfully forecasted with several experimental tests from different sectors.

1. Introduction

The price variation of stock market is a dynamic system and the chaotic behavior of the stock price movement duplicates complication of the price prediction; however, the highly non-linear, dynamic complicated domain knowledge inherent in the stock market makes it very difficult for investors to make the right investment decisions promptly. It is necessary to develop an intelligent system to get real-time pricing information, reduce one obsession of investors and help them to maximize their profits.

There are two common approaches has used for market analysis which are fundamental analysis and technical analysis. A fundamental analysis involved some statistics of the macroeconomics data as well as the basic financial status of the company. After taking all these factors into account, the analyst will then make a decision of selling or buying a stock. Another approach is based on the historical financial time series data called technical analysis. However, financial time series show quite complicated patterns (for example, trends, abrupt changes, and volatility clustering) and such series are often non-stationary, whereby a variable has no clear tendency to move to a fixed value or a linear Trend (Chang & Liu, 2008).

Stock price prediction has always been a subject of interest for most investors and professional analysts. Nevertheless, finding out the best time to buy or to sell has remained a very difficult task because commission (remuneration for services rendered) is commonly overlooked when doing research relating to stock market prediction; however, if any model is actually implemented it is going to incur fees which could greatly affect the profit predicted by the model. Chen and Linkens (2004) considers three different levels of commissions and how it would affect the best buying strategy used by Investors. Tehran Stock Exchange has evolved into an exciting and growing marketplace where individual and institutional investor trade securities of over 420 companies. TEPIX is a weighted market value all share prices appearing on the TSE Price Board. TEPIX calculation method is as follows:

$$TEPIX = \frac{\sum P_{it}C_{it}}{\sum P_{ib}C_{it}}(Base - Value),$$

where P_{it} and P_{ib} represent share price of company i, respectively at time t and at the close of trading on March 21st, 1990; and C shows the total number of shares. And other important index in this market contain Industrial Index, Financial Index, Top 50 Companies Index. The remaining sections of this paper are organized as follows. Section 2 provides the literature review. Section 3 introduces the essential concepts of a type of fuzzy rule based systems called TSK and in its sub section issue of input variable and performance measurement has been considered. Section 4 describes ANFIS struc-

* Corresponding author. Tel.: +98 21 64542782.
E-mail addresses: Esfahaa@aut.ac.ir (A. Esfahanipour), Werya_Aghamiri@aut.ac.ir (W. Aghamiri).

0957-4174/$ - see front matter © 2009 Published by Elsevier Ltd.
doi:10.1016/j.eswa.2009.11.020

ture. Section 5 discusses the Fuzzy c-mean clustering algorithms. Section 6 presents the underlying research methodology. In Section 7 we can see experimentation design and results and Comparisons of different forecasting models. Finally, concluding remarks are given in Section 8.

2. Literature survey

It is too difficult to forecast stock price variation and the price fluctuation behaves more like a random walk and time varying. In a stock market, how to forecast stock prices accurately and find the right time to trade are obviously of great interest to investors. To reach this goal, two approaches can be employed: statistical and Artificial Intelligence (AI).

The statistics school includes autoregressive integrated moving average (ARIMA), generalized autoregressive conditional heteroskedasticity (GARCH) volatility (Franses & Ghijsels, 1999), and smooth transition autoregressive (STAR) (Sarantis, 2001). These models depend on the assumption of linearity among variables and normal distribution.

However, the assumption of linearity and normal distribution may not hold even though it has been shown successful in dealing with stock price movement in past decades. On the other hand, with greater in the stock market and the increasing need for more efficient forecasting models, the AI school, operating without the limitation of such an assumption, outperforms the conventional statistical methods experimentally (Enke & Thawornwong, 2005; Hansen & Nelson, 2002; Ture & Kurt, 2006; Zhang, 2003). During the last decade, stocks and future traders have come to rely upon various types of intelligent systems to make trading decisions. Recently, artificial neural networks (ANNS) have been applied to this area (Aiken & Bsat, 1999; Chang, Wang, & Yang, 2004; Chi, Chen, & Cheng, 1999; Kimoto & Asakawa, 1990; Lee, 2001; Yao & Poh, 1995; Yoon & Swales, 1991). These models, however, have their limitations because of the great noise and complex dimensionality of stock price data and besides, the quantity of data itself and the input variables may also middle in each other. Therefore, the result may not be very satisfactory.

Other soft computing methods are also applied in the prediction of stock price and these Soft Computing (SC) approaches are supposed to use quantitative inputs, like technical indices, and qualitative factors, like political effects, to automate stock market forecasting and trend analysis. Kuo, Chen, and Hwang (2001) uses a genetic algorithm base fuzzy neural network to measure the qualitative effects on the stock price. Variable selection is critical to the success of any network for the financial viability of a company. They applied their system to the Taiwan stock market. Aiken and Bsat (1999) use a FFNN (Feed Forward Neural Network) trained by a genetic algorithm (GA) to forecast three-month US Treasury Bill rates. They come to this conclusion that an NN can accurately predict these rates. Thammano (1999) used a neuro-fuzzy model to predict future values of Thailand's largest government-owned bank. He concluded that the neuro-fuzzy architecture was able to recognize the general characteristics of the stock market faster and more accurately than the basic back-propagation algorithm. Baba, Inoue, and Asakawa (2000) used NNs and GAs to construct an intelligent decision support system (DSS) for analyzing the Tokyo Stock Exchange Prices Indexes (TOPIX). The essential feature of their DSS was that it projected the high and low TOPIX values four weeks into the future and suggested buy and sell decisions based on the average projected value and the then-current value of the TOPIX. Kim and Han (2000) used a NN modified by a GA to predict the stock price index. In this instance, the GA was used to reduce the complexity of the feature space, by optimizing the thresholds for feature discretization, and to optimize the connection weights between layers. They concluded that the GA approach outperformed the conventional models. Abraham, Philip, and Saratchandran (2003) investigate how the seemingly chaotic behavior of stock markets could be well represented using several connectionist paradigms and soft computing techniques. To demonstrate the proposed technique, they analyzed the 7 year's Nasdaq-100 main index and 4 year's NIFTY index values. They concluded that all the connectionist paradigms considered could represent the stock indices behavior very accurately. Recently, therefore regarding to neural network's ability to model any given function and in other wise, fuzzy rule based systems mimics the crucial ability of the human mind to summarize data and focus on decision-relevant information, retaining good results by the proposed model is our expectation in this area.

3. A TSK type fuzzy rule based system

This research is aimed to predict the future price of a stock by the technical index input to the adapted TSK fuzzy model. For a first order Takagi–Sugeno model, a common rule is represented as follows: If x_1 is A_1, and x_2 is A_2, then $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2$ where x_1 and x_2 are linguistic variables and A_1 and A_2 are corresponding fuzzy sets and β_0, β_1, β_2 are linear parameters. Usually determining the linear parameters by using the least-squares algorithm and the membership function parameters are fine tuned using a neural network learning method. Initial rules are generated using the grid partitioning method. The various rule parameters during the TSK adaption process into ANFIS will be fine tuned. This research mainly studies the fluctuations of short-term stock prices and tries to develop a forecasting model using TSK type fuzzy rule based approach. The TSE index. are selected for studying purposes, and for more evaluation of the proposed model we test it on Tehran stock market index.

3.1. Input variables

There are so many factors that might effect stock prices. Many papers have dealt with input selection to mapping financial indexes and stocks. We can divide Inputs into two different types of inputs, financial and political (which tend to be qualitative). Quah and Srinivasan (1999) specified 5 key factors which will influence the stock price movement, that is yield, liquidity, risk, growth, and momentum. Izumi and Ueda (1999) stated that stock returns can be affected directly by macroeconomic factors such as inflation and short-term interest rate. As for the measure of system performance, Yao and Poh (1995) indicated an example that a model with a low normalized mean square error (NMSE) had a lower return than a model with a higher NMSE. In order to enhance the understanding of inexperienced traders and other people, Brownstone (1996) suggested using percentages to measure performance. Chen, Leung, and Daouk (2003) used a sliding window to forecast the next day's price of the index. The network was retrained with the most recent 68 days of input attempting to forecast the next day. In other wise to show the current tendency of the stock price fluctuations, technical indexes are calculated from the stock price's time series, trading volumes and time which are following a set of formula. We can apply these indexes for decision making to evaluate oversold or overbought in the stock market. Basically, the technical index can be classified as index for TSE movement or particular stock price variations, such as KD, RSI, MACD, MA, BIAS. As Chang et al. (2004) says, seven technical indexes are illustrated as shown in Table 3.1. Because ease of compute and further more for fair comparison with other model has used these factors as input variable for the model.

Table 3.1
Technical indexes used as input variables.

Technical index	Explanation
Six days moving average (MA)	Moving averages are used to emphasize the direction of a trend and smooth out price and volume fluctuations that can confuse interpretation
Six days bias (BIAS)	The difference between the closing value and moving average line, which uses the stock price nature of returning back to average price to analyze the stock market
Six days relative strength index (RSI)	RSI compares the magnitude of recent gains to recent losses in an attempt to determine overbought and oversold conditions of an asset
Nine days stochastic line (K,D)	The stochastic line K and line D are used to determine the signals of over-purchasing, overselling, or deviation
Moving average convergence and divergence (MACD)	MACD shows the difference between a fast and slow exponential moving average (EMA) of closing prices. Fast means a short-period average, and slow means a long period one
13 days psychological line (PSY)	PSY is the ratio of the number of rising periods over the total number of periods. It reflects the buying power in relation to the selling power
Volume	Volume is a basic yet very important element of market timing strategy; volume provides clues as to the intensity of a given price move

3.2. Performance measure

Performance measurement has been used for evaluating performance and error evaluation of results of model. Using these measures widely for assessing gap between real output and predicted results gained by model.

The literature review show us several PM such as MSE (Mean Squared Error), RMSE (Root Mean Square Error), NMSE (Normalized Mean Squared Error), MAE (Mean Absolute Error), MAPE (Mean Absolute Percent Error),

MAPE is applied in this research as an evaluation basis. Its formula is presented as follows:

$$\mathrm{MAPE} = \frac{100}{P} \times \sum_{p=1}^{P} \left| \frac{d_p - z_p}{d + p} \right|,$$

where z: forecasted output; d: actual output; P: total number of records.

4. ANFIS structure

Jang (1993, 1995, 1997) developed ANFIS which is a neuro-fuzzy system. A feed-forward neural network structure has been used for this aim where each layer is a neuro-fuzzy system component (Fig. 4.1). It simulates TSK (Takagi–Sugeno–Kang) fuzzy rule (Sugeno & Kang, 1988) of type-3 where the consequent part of the rule is a linear combination of input variables and a constant. The final output of the system is the weighted average of each rule's output. The form of the type-3 rule simulated in the system is as follows:

$$\text{IF } x_1 \text{ is } A_1 \text{ AND } x_2 \text{ is } A_2 \text{ AND}, \ldots, \text{AND } x_p \text{ is } A_p$$
$$\text{THEN } y = c_0 + c_1 x_1 + c_2 x_2 + \cdots + c_p x_p$$

where x_1 and x_2 are the input variables, A_1 and A_2 are the membership functions, y is the output variable, and c_0, c_1, and c_2 are the consequent parameters.

The neural network structure includes of six layers.

- Layer 0 is the input layer. It has n nodes here: n is the number of inputs to the system.
- The fuzzy part of ANFIS is mathematically incorporated in the form of membership functions (MFs).
  A membership function $\mu_{A_i}(x)$ is any continuous and piecewise differentiable function which transforms the input value x into a membership degree, which means a value between 0 and 1. The most widely applied membership function is the generalized bell (gbell MF), which is described by the three parameters, a, b, and c (Eq. (4.1)). Therefore, Layer 1 is the fuzzification layer in which each node represents a membership value to a linguistic term as a Gaussian function with the mean;

$$\mu_{A_i}(x) = \frac{1}{1 + \left[\left(\frac{x - c_i}{a_i} \right) \right]^{b_i}}, \tag{4.1}$$

where a_i, b_i and c_i are parameters of the function. These are adaptive parameters. During the learning stage the back-propagation algorithm adapts their values. Changes in the membership function of the linguistic term, A_i depends on the changes of the values of the parameters. These parameters are called premise parameters. In that layer there exist $n \times p$ nodes where n is the number of input variables and p is the

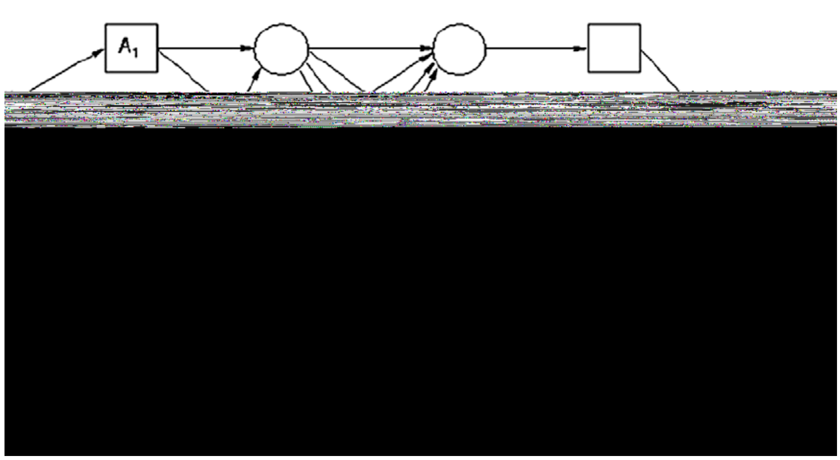

Fig. 4.1. Basic ANFIS structure.

number of membership functions. For example, if size is an input variable and there existWe can create two nodes in the first layer for two linguistic values of size that are SMALL and LARGE and the membership values of input variable size are indicated by them to the linguistic values SMALL and LARGE.

- Each node in Layer 2 provides the strength of the rule by means of multiplication operator. It performs AND operation.

  $$W_i = \mu_{A_i} \times \mu_{B_i}.$$

  The multiplication of the input values is calculated by means of every node in this layer and gives the product as the output like what is shown in the above equation. The membership values represented by $\mu_{A_i}(x_0)$ and $\mu_{B_i}(x_1)$ are multiplied in order to find the firing strength of a rule where the variable x_0 has linguistic value *Ai* and x_1 has linguistic value *Bi* in the antecedent part of Rule I.
  There are p^n nodes indicating the number of rules in Layer 2. Each node is representative of the antecedent part of the rule. We will have four rules in the system antecedent parts, when two variables exist in the system which are x_1 and x_2 that show two fuzzy linguistic values, SMALL and LARGE, whose are as follows:

  IF x_1 is SMALL AND x_2 is SMALL
  IF x_1 is SMALL AND x_2 is LARGE
  IF x_1 is LARGE AND x_2 is SMALL
  IF x_1 is LARGE AND x_2 is LARGE

- Layer 3 is the normalization layer which normalizes the strength of all rules according to the equation.

  $$\bar{w}_i = \frac{w_i}{\sum_{j=1}^{R} w_j},$$

  where w_i is the firing strength of the *i*th rule which is computed in Layer 2. Node *i* computes the ratio of the *i*th rule's firing strength to the sum of all rules' firing strengths. There are *pn* nodes in this layer.
- Layer 4 is a layer of adaptive nodes. Every node in this layer computes a linear function where the function coefficients are adapted by using the error function of the multi-layer feed-forward neural network.

  $$\bar{w}_i f_i = \bar{w}_i(p_0 x_0 + p_1 x_1 + p_2),$$

  p_i s are the parameters where $i = n + 1$ and n is the number of inputs to the system (i.e., number of nodes in Layer 0). In this example, since there exist two variables (x_1 and x_2), there are three parameters p_0, p_1 and p_2) in Layer 4 and w_i is the output of Layer 3. The parameters are updated by a learning step. Kalman filtering based on least-squares approximation and back-propagation algorithm is used as the learning step.
- Layer 5 is the output layer whose function is the summation of the net outputs of the nodes in Layer 4. The output is computed as:

  $$\sum \bar{w}_i f_i = \frac{\sum_i w_i f_i}{\sum_i w_i},$$

  where $w_i f_i$ is the output of node *i* in Layer 4. It denotes the consequent part of rule *i*. The overall output of the neuro-fuzzy system is the summation of the rule consequences. ANFIS uses a hybrid learning algorithm in order to train the network. For the parameters in the layer 1, back-propagation algorithm is used. For training the parameters in the Layer 4, a variation of least-squares approximation or back-propagation algorithm is used.

5. The Fuzzy C-Means clustering algorithm

One of the common unsupervised clustering algorithm used in the literature is Fuzzy C-Means (FCM). The original algorithm was described first by Bezdek (1973), derivatives have been described with modified definitions for the norm and prototypes for the cluster centroids (Kim, Lee, & Lee, 2004; Man & Gath, 1994). FCM tries to find the most particular point in each cluster, which can be regarded as the "centroid" of the cluster and, then, the grade of membership for each object in the clusters. Such purpose is attained by minimizing the objective function. A commonly used objective function is membership weighted within cluster error defined as follow:

$$\textit{Minimize } J_m(U, V) = \sum_{j=1}^{n} \sum_{i=1}^{c} u_{ij}^m \|x_j - v_i\|^2, \tag{5.1}$$

where n is the total number of patterns in a given data set and c is the number of clusters; $X = x_1, x_2, \ldots, x_n \subset R_s$ and $V = v_1, \ldots, v_c \subset R_s$ are the feature data and cluster centroids; and $U = [u_{ij}]_{c\times n}$ is a fuzzy partition matrix composed of the membership grade of pattern x_j to each cluster *i*. $||xj - vi||$ is the Euclidean norm between *xj* and *vj*. The weighting exponent *m* is called the fuzzifier which can have influence on the clustering performance of FCM. The constrained optimization problem can be solved by the cluster centroids and the respective membership functions in (5.1) are given by the following equations:

$$V_j = \frac{\sum_{j=1}^{n} (u_{ij})^m x_j}{\sum_{j=1}^{n} (u_{ij})^m}, \quad 1 \leqslant i \leqslant c \tag{5.2}$$

$$U_{ij} = \left[\sum_{k=1}^{c} \left(\frac{\|x_j - v_i\|^2}{\|x_j - v_k\|^2} \right)^{1/(m-1)} \right]^{-1}, \quad 1 \leqslant i \leqslant c, \quad 1 \leqslant j \leqslant n. \tag{5.3}$$

Eqs. (5.2) and (5.3) consist of an iterative optimization procedure. The goal is to iteratively improve a sequence of sets of fuzzy clusters until no further improvement in $J_m(U, V)$ is possible. The FCM algorithm is executed in the following steps:

Step 1: Given a pre-selected number of cluster *c*, *a* chosen value of *m*, initialize memberships *uij* of x_j belonging to cluster *i* such that

$$\sum_{i=1}^{c} u_{ij} = 1. \tag{5.4}$$

Step 2: Calculate the fuzzy cluster centroid98.6645T4158.1853Tf5.4238

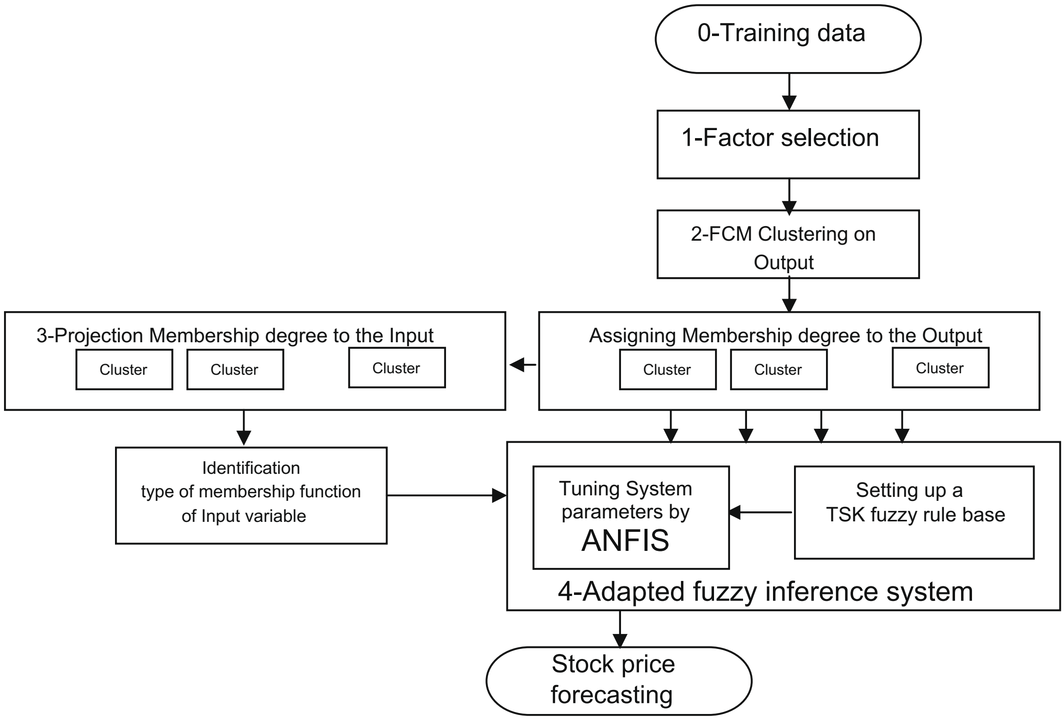

Fig. 6.1. Framework of the proposed model.

Step 3: Projection of membership values of the output onto the input spaces to generate the membership values of the inputs.

Step 4: Tuning the parameters of membership function of input and output variables by using adaptive nero-fuzzy inference system.

6.1. Factor selection by stepwise regression

Seven technique indices are suggested as the independent variables According to the numbers gained and indices selected in Section 3.1. The research adopts stepwise regression to analyze and select variables, Stepwise regression is a popular and extremely effective method for establishing regression models. And further simplifies it to improve the forecasting accuracy in order to avoid the interferences made by indices in the training process with low impacts themselves or overshadowed to the model affected by mutual function between two indices, which may decrease the explanation ability of such model and increase the error occurrence. This algorithm is used to sort variables out and leave more effective ones in the model. The most widespread stepwise regression process works as follows. The user first identifies the response, y, and the set of potentially important independent variables, x_1, $x_2, \ldots, x_k$, where k is generally large. The response and independent variables are then entered into the regression process (the stepwise process begins) (McClave, Benson, & Sincich (2005)).

The algorithm of stepwise regression constructs the model via a series of iterations. Each iteration involves one of two processes, namely either adding a variable to the model (referred to here as the selection process) or removing a variable from the model (elimination). The process of stepwise regression involves of these steps as follows:

1. Consider all possible regressions using one explanatory variable. Select that variable with the largest t-ratio. If the t-ratio is not significance, then do not select any variables and halt the process.
2. The next variable to enter is the one that makes the most significant contribution. To enter, the t-ratio must exceed a specified t-value.
3. Next, delete the variable that makes the smallest contribution. The deleted t-ratio must be below a specified t-value.
4. Repeat steps 2 and 3 until all possible additions and deletions have been performed.

When only one variable is being considered recall that (t-ratio) 2 = F-ratio. SPSS is used for variable selection and setting up the regression forecasting model.

6.2. Clustering the output space and determination of the number of rules

In this paper, first, the output data is clustered and the primary membership grades of the output clusters are generated. For this purpose, Sugeno and Yasukawa (1993) method is used. We first partition the output space and then obtain the input space clusters by "projecting" the output space partition onto each input variable space, separately.

In order to carry out the process of encoding the output space, we consider one of the most applicable fuzzy clustering algorithms, i.e., Fuzzy C-Mean clustering.

It is required to obtain a cluster validity criterion in order to determine the optimal number of clusters presented in the data. In the case of the FCM algorithm, there are many validation indices found in the literature (Baduska, 1995; Gath & Geva, 1989) can be divided into two well-known categories first validity indices involving only the membership values such as V_{PE}, V_{PC}, V_{WPE}, V_{MPC}, V_{KYI}, and V_P and validity indices involving the membership values and the data set such as V_{FS}, V_{XB}, V_K, V_T, and V_{SC}.

In this research, Euclidean Distance has been selected as similarity measure of the set of data.

- The proposed cluster validity based on similarity measure.

Cluster analysis aims at identifying groups of similar objects, therefore helps to discover distribution of patterns and interesting correlations in large data sets. However, most clustering algorithms need to know the number of classes to look for. This is an unsupervised method and in most cases, user will not have any prior knowledge about the number of classes, we are separating the data set into, is larger or smaller than the actual number of classes. If it is larger then one or more good compact clusters may be broken. If it is smaller then more than one separate cluster may be merged.

Thus finding the right number of clusters is an important problem. The problem for finding an optimal c is usually called cluster validity. Once the partition is obtained by a clustering method, the validity function can help us to validate whether it accurately presents the structure of the data set or not.

According to the issues has been marked in the last paragraph a number of validity indices for fuzzy clustering exist in the literature. Early indices such as the partition coefficient and classification entropy make use only of membership values and have the advantage of being easy to compute.

(a) Bezdek attempted to define a performance measure based on minimizing the overall content of pairwise fuzzy intersection in U, the partition matrix. He proposed cluster validity index for fuzzy clustering (Bezdek, 1981; Trauwaert, 1988): partition coefficient (*PC*). The index was defined as

$$V_{PC} = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}^2.$$

The *PC* index indicates the average relative amount of membership sharing done between pairs of fuzzy subsets in U, by combining into a single number, the average contents of pairs of fuzzy algebraic products. The index values range in $[1/c, 1]$, where c is the number of clusters. In general, we find an optimal cluster number c^* by solving $\max V_{PC}$, $2 \leqslant c \leqslant n-1$ to produce the best clustering performance for the data set X.

(b) Bezdek proposed the partition entropy (*PE*) (Bezdek, 1974, 1981) was defined as

$$V_PE = -\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n}u_{ij}\log_a u_{ij},$$

where a is the base of the logarithm. The *PE* index is a scalar measure of the amount of fuzziness in a given U. According to Bezdek (1981), the limitation of the *PE* can be attributed to its apparent monotonicity and to an extent, to the heuristic nature of the rationale underlying its formulation. The index is computed for values of c greater than 1 and its values range in $[0, \log_a c]$. In general, we find an optimal c^* by solving $\min V_{PE}$, $2 \leqslant c \leqslant n-1$ to produce the best clustering performance for the data set X.

6.3. Projection of membership functions of output onto input spaces

After selection of the significant input variables, suitable membership functions should be determined for them. With respect to the proposed approach of Fazel Zarandi (1998), first, the ranges in which input variable membership functions that adapt value 1 are determined. Then, the data points are classified, using FCM by given m and c, which were determined in the pervious stage.

6.4. Parameters identification of membership function of input and output variables by ANFIS

After determining the number of rules and the type of the membership function of the variables in antecedent and consequent part in the previous stages, now we can establish structure of the ANFIS such as number of the nodes in the layers and type of transfer function in the nodes.

7. Experimental setup and results

Regarding to make comparison between proposed model and other models have experienced on Taiwan Stock Exchange index we use the data set including TSE index has been decomposed into three different sets: the training data, test data and validation data. The data for TSE index are from July 18, 2003 to December 31, 2005 totally 614 records and the first 494 records will be training data and the rest of the data, i.e., 120 records will be test data at the same data combination of the TEPIX has used from April 20, 2006 to January 31, 2009 contain 479 records. Between the factors, we will test each factor using stepwise regression analysis and identify the factor that will affect the final forecasted results significantly. The final combination of the factors will be finalized after the analysis. The factors selected finally are MA6 and BIAS6 these two index and the output variables are TSE index.

7.1. Forecasted results in TSE index and Some Tehran Stock Exchange Indexes

According to the MAPE value as a performance measurement, accurate results gained by proposed model close to 98.7% in TSE and 97.3% in TEPIX. The comparisons of different models such as BPN (Back Propagation Neural networks), multiple regression, (respectively, MATLAB and SPSS is used for implementation) the TSK fuzzy rule model proposed by Chang and Liu (2008) and our proposed model are listed in Table 7.1. The forecasted results from ANFIS based model are much better than those from BPN, multiple regression or TSK fuzzy rule model tuned by SA, which justify the ANFIS based model is the best.

Tehran Stock Exchange Index data has been used in the model by more several experimental test and results have been presented in Table 7.2.

Table 7.1
Comparison of different methods.

	Methods			
	Proposed model (9 rules)	BPN	Multiple regression analysis	TSK Tuned by SA (7 rules)
Data MAPE of Taiwan stockExchange index (TSE)	1.3%	4.29%	7.08%	2.4%

Table 7.2
Results of the model on the different Tehran Stock Exchange Indexes.

	Indexes			
	TEPIX	Index of top 50 Companies	Industry Index	Index of Financial Group
MAPE of proposed model	2.4%	1.85%	2.02%	1.03%

8. Conclusion

Proposed model contain a TSK fuzzy rule based system which apply significant technical indexes as input variable. Significant variable gained during variable selection process; that research method has used stepwise regression as well-defined algorithm in this area. In other side consequent part of the TSK approximated by linear combination of the technical indexes offered by factor selection stage. Number of optimal rules defined by FCM algorithm and further more membership degree of the input variables of each rules are assigned during projection of membership degree of the desired stock price (close price the next day) into its variable in premise. In other hand initial membership function of the premise variable approximately defined as Gaussian function. Adaptive neuro-fuzzy system is applied to further fine-tune the parameters of the linear combination of the input variables. Therefore, the forecasting capability of the system is greatly improved. Finally, the system is tested on the TSE and some Tehran Stock Exchange Indexes such as Total Index, Top 50 company Index, Industry Index, Financial Group Index. And all the performance results are outperforming other approaches such as BPN, or multiple regression analysis.

This research is just a beginning and the long term goal is to predict the trend of the price variation by including various influential factors such as macro economic change, political reasons, fundamental analysis and the technical index etc. As a result, the system can be further applied for the daily trading purpose.

References

Abraham, A., Philip, N. S., & Saratchandran, P. (2003). Modeling chaotic behavior of stock indices using intelligent paradigms. *Neural, Parallel and Scientific Computations, 11*, 143–160.

Aiken, M., & Bsat, M. (1999). Forecasting market trends with neural networks. *Information Systems Management, 16*(4), 42–48.

Baba, N., Inoue, N., & Asakawa, H. (2000). Utilization of neural networks and GAs for constructing reliable decision support systems to deal stocks. In *IEEE–INNS–ENNS international joint conference on neural networks (IJCNN'00)* (Vol. 5, pp. 5111–5116).

Baduska, R. (1995). New approach to constructing fuzzy relational models from data. In *Proceedings of the European control conference* (pp. 1207–1212), Rome.

Bezdek, J. C. (1973). Fuzzy mathematics in pattern classification. Ph.D. dissertation, Cornell University, Ithaca, NY.

Bezdek, J. C. (1974). Cluster validity with fuzzy sets. *Journal of Cybernetics, 3*, 58–73.

Bezdek, J. C. (1981). *Pattern recognition with fuzzy objective function algorithms*. New York: Plenum Press.

Brownstone, D. (1996). Using percentage accuracy to measure neural network predictions in stock market movements. *Neurocomputing, 10*, 237–250.

Chang, P. C., & Liu, C. H. (2008). A TSK type fuzzy rule based system for stock price prediction. *Expert Systems with Applications, 34*, 135–144.

Chang, P. C., Wang, Y. W., & Yang, W. N. (2004). An investigation of the hybrid forecasting models for stock price variation in Taiwan. *Journal of the Chinese Institute of Industrial Engineering, 21*(4), 358–368.

Chen, A. S., Leung, M. T., & Daouk, H. (2003). Application of neural networks to an emerging financial market: Forecasting and trading the Taiwan stock index. *Computers and Operations Research, 30*, 901–923.

Chen, M. Y., & Linkens, D. A. (2004). Rule-base self-generation and simplification for data-driven fuzzy models. *Fuzzy Sets and Systems, 142*, 243–265.

Chi, S. C., Chen, H. P., & Cheng, C. H. (1999). A forecasting approach for stock index future using grey theory and neural networks. In *IEEE international joint conference on neural networks* (pp. 3850–3855).

Enke, D., & Thawornwong, S. (2005). The use of data mining and neural networks for forecasting stock market returns. *Expert Systems with Applications, 29*, 927–940.

Fazel Zarandi, M. H. (1998). Aggregate system analysis for prediction of tardiness and mixed zones of continuous casting with fuzzy methodology. Ph.D. thesis.

Franses, P. H., & Ghijsels, H. (1999). Additive outliers, GARCH and forecasting volatility. *International Journal of Forecasting, 15*(1), 1–9.

Gath, I., & Geva, A. B. (1989). Unsupervised optimal fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligent, 11*(7), 773–781.

Hansen, J. V., & Nelson, R. D. (2002). Data mining of time series using stacked generalizers. *Neurocomputing, 43*(1–4), 173–184.

Izumi, K., & Ueda, K. (1999). Analysis of exchange rate scenarios using an artificial market approach. In *Proceedings of the international conference on artificial intelligence* (Vol. 2, pp. 360–366).

Jang, J.-S. R. (1993). *IEEE Transactions on Systems Man and Cybernetics, 23*, 665.

Jang, J.-S. R. (1997). *Neuro-fuzzy and soft computing*. New Jersey: Prentice-Hall.

Kim, K. J., & Han, I. (2000). Genetic algorithms approach to feature discretization in artificial neural networks for the prediction of stock price index. *Expert Systems with Applications, 19*, 125–132.

Kim, D. W., Lee, K. H., & Lee, D. (2004). On cluster validity index for estimation of the optimal number of fuzzy clusters. *Pattern Recognition, 37*, 2009–2025.

Kimoto, T., & Asakawa, K. (1990). Stock market prediction system with modular neural network. In *International joint conference on neural network* (Vol. 77, pp. 1–6).

Kuo, R. J., Chen, C. H., & Hwang, Y. C. (2001). An intelligent stock trading decision support system through integration of genetic algorithm based fuzzy neural network and artificial neural network. *Fuzzy Sets and Systems, 118*, 21–24.

Lee, J. W. (2001). Stock price prediction using reinforcement learning. In *International joint conference on neural networks* (Vol. 44, pp. 690–695).

Man, Y., & Gath, I. (1994). Detection and separation of ring-shaped clusters using fuzzy clustering. *IEEE Transactions on Pattern Analysis and Machine Intelligent, 16*(8), 855–861.

McClave, J. T., Benson, P. G., & Sincich, T. (2005). *Statistics for business and economics* (9th ed.). Prentice-Hall.

Quah, T. S., & Srinivasan, B. (1999). Improving returns on stock investment through neural network selection. *Expert Systems with Applications, 17*, 295–301.

Sarantis, N. (2001). Nonlinearities, cyclical behavior and predictability in stock markets: International evidence. *International Journal of Forecasting, 17*(3), 459–482.

Sugeno, M., & Kang, G. T. (1988). *Fuzzy Sets and Systems, 28*, 15.

Sugeno, M., & Yasukawa, T. (1993). A fuzzy logic based approach to qualitative modeling. *IEEE Transactions on Fuzzy Systems, 1*, 7–31.

Thammano, A. (1999). Neuro-fuzzy model for stock market prediction. In C. H. Dagli, A. L. Buczak, J. Ghosh, M. J. Embrechts, & O. Ersoy (Eds.) Smart engineering system design: Neural networks, fuzzy logic evolutionary programming, data mining, and complex systems. In *Proceedings of the artificial neural networks in engineering conference (ANNIE '99)* (pp. 587–591). New York: ASME Press.

Trauwaert, E. (1988). On the meaning of Dunn's partition coefficient for fuzzy clusters. *Fuzzy Sets and Systems, 25*, 217–242.

Ture, M., & Kurt, I. (2006). Comparison of four different time series methods to forecast hepatitis A virus infection. *Expert Systems with Applications, 31*, 41–46.

Yao, J., & Poh, H. L. (1995). Forecasting the KLSE index using neural networks. In *International conference on neural networks* (Vol. 2, pp. 1012–1017).

Yoon, Y., & Swales, J. (1991). Prediction stock price performance: a neural network approach. In *Proceeding of the twenty-fourth annual Hawaii international conference on system science* (pp. 156–162).

Zhang, G. (2003). Time series forecasting using a hybrid ARIMA and neural network model. *Neurocomputing, 50*, 159–175.