Research Journal of Applied Sciences 11 (10): 1069-1075, 2016
ISSN: 1815-932X

Combining Fastica with Back Propagation Algorithms for the Forecasting of Gulf Cooperation Council Stock Market

[1]Yasen Rajihy and [2]Fadheela Sabri Abu-Almash
[1]College of Information Technology, University of Babylon, Babylon, Iraq
[2]Ministry of Higher Education and Scientific Research, Baghdad, Iraq

Abstract: In this study we present a best model of forecasting the Gulf Cooperation Council (GCC) stock market by combining FastICA, TnA and BP algorithms (FastICA-BP) model. In this model, FastICA is firstly used to analyse the raw dataset to get components which are independent of each other. Secondly, TnA approach applied to identify and remove the IC representing the noise. And finally, BP algorithm used to predict the stock market using the filtered components. To evaluate the performance of the proposed model, Al Rajihy Islamic bank used as illustrative example in this study. The experimental results show that the proposed model outperforms the BP model using original data set (Model I) and BP model using non-filtered dataset (Model II).

Keywords: Independent component analysis, FastICA, back propagation, stock market forecasting, algorithms

INTRODUCTION

Forecasting in stock market plays a key role in the theory and practice of investing, especially given the progress in automation of turnover on capital markets which trading systems use for generation of purchase and sale signals. Various prediction techniques involve implementing complex procedures that employ statistical, technical and fundamental analysis. The categories of such solutions are actually failed because the stock market must be treated as an enormously complex system that changes, sometimes dramatically over time and is influenced by numerous factors that are often mutually correlated and contain feedback, thus having strongly nonlinear characteristics in their linkages. The variability of capital markets over time and those nonlinear linkages combine to demonstrate the need for continual improvement of models of stock price shaping on the stock market. The promising approach in solving problems of such type is the application of Artificial Neural Networks (ANN). ANN are considered more effective in forecasting than other methods because they can learn and recognise patterns and relationships between data objects (Liu and Wang, 2011).

Independent Component Analysis (ICA) is the next method with real promise in this area. This is a modern technique of signal processing well known in the field of data pattern recognition. ICA is a mathematically grounded, quantitative method that discovers the latent factors that underlie a set of apparently random signals (Hyvarinen, 1997, 1999).

ICA is primarily used in applications for the analysis of input data. One of the major challenges in stock exchanges is the amount of noise in financial data that reduces prediction efficiency and can lead to problems in under or over estimation. The detection and elimination of noise is thus a very important issue and at the same time, poses a serious challenge in building a prediction model (Lu *et al*., 2009). Testing-and-Acceptance (TnA) is regarded one of the most popular technique to solve this problem by using Relative Hamming Distance (RHD) reconstruction error.

In this study we present a best model of forecasting the Gulf Cooperation Council (GCC) stock market by combining FastICA, TnA and BP techniques (FastICA-BP) model. In this model, FastICA is firstly used to analyse the raw dataset to get components which are independent of each other. Secondly, TnA approach has been applied to identify and remove the IC representing the noise. And finally, BP algorithm has been used to predict the stock market using the filtered components.

One of the most stock indexes in GCC, namely, Al Rajhi Islamic bank has been used as illustrative example to evaluate the performance of the proposed model. The experimental results show that the proposed model outperforms the BP model using original data set (Model I) and BP model using non-filtered data set (Model II).

MATERIALS AND METHODS

Independent component analysis: Independent component analysis (ICA) is a modern signal processing

Corresponding Author: Yasen Rajihy, College of Information Technology, University of Babylon, Babylon, Iraq

technique. It is a well-known mathematical method for revealing latent components that underlie sets of random signals (variables) and identify hidden structures in high dimensional data (Hyvarinen and Oja, 2000). ICA can be considered an improvement over Principal Component Analysis (PCA) and Factor Analysis (FA).

The ICA technique was developed at the beginning of the 1980s and found its first application in the field of neural network modeling (Hyviirinen *et al*., 2001). During that decade, it remained mostly unknown at the international level. Now ICA methods are widely applied to different application fields. It has become a well-known technique in fields such as neural networks, signal processing and advanced statistics.

The main mathematical problem of ICA can be described as follows: We observe a random vector $x = (x_1, x_2,...., x_m)$ of m dimension that is modeled as linear combinations of n dimension random vector $s = (s_1, s_2,...., s_n)$:

$$\begin{aligned} x_1 &= \alpha_{11}s_1 + \alpha_{12}s_2 + ... + \alpha_{1n}s_n \\ x_2 &= \alpha_{21}s_1 + \alpha_{22}s_2 + ... + \alpha_{2n}s_n \\ x_m &= a_{m1}s_1 + a_{m2}s_2 + ... + a_{mn}s_n \end{aligned} \quad (1)$$

where, n are some real mixing coefficients. The components of s are supposed to be statistically mutually independent. The matrix representation of Eq. 1 can be expressed as:

$$\begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix} = \begin{pmatrix} \alpha_{11}\alpha_{12}...\alpha_{1n} \\ \alpha_{21}\alpha_{22}...\alpha_{2n} \\ \vdots \\ \alpha_{m1}\alpha_{m2}...\alpha_{mn} \end{pmatrix} \begin{pmatrix} s_1 \\ s_2 \\ \vdots \\ s_n \end{pmatrix} \quad (2)$$

or simply:

$$x = As \quad (3)$$

where, A is unknown m×n constant matrix should be evaluated (estimated). The basic ICA model states that observed data are produced by mixing the ICs (s_j). These ICs are the hidden variables, i.e., they are not observable directly. Also the mixing coefficients α_{ij} are unknown. ICA solely uses the observed data x_i to estimate both s_j and the mixing coefficients α_{ij}.

The task in ICA is to find both the latent variables or sources s_j and the mixing process, i.e., finding the mixing matrix in A a linear case. One popular solution finds a demixing matrix W so that variables y_j in $y = Wx$ are estimates of s_j up to scaling and permutation. Such a matrix can be treated as an approximation of the W inverse matrix of A, up to scaling and permutation of the rows of W.

Often, the latent variables s_j are estimated one by one by finding a column vector W_j (this will be stored as a row of W) such that $y_j = w_j^T x$ is an estimate of s_j (Bingham, 2003). In order to estimate a demixing matrix W (i.e. to estimate ICs), numerous algorithms have been developed with various approaches. FastICA is regarded as one of the best algorithms to solve the ICA model (Hyvarinen and Oja, 2000).

Back propagation artificial neural network: Artificial Neural Network (ANN) can be applied to a wide variety of problems such as image recognition, classifying patterns, signal processing, speech production, medicine and optimization problems. Recently, applications of ANN have been increasing in business, informational technology, environmental area, politics and economics. ANN are suitable methods to analyze economical time series and issues related to stock markets (Khaze *et al*., 2013). Many studies on stock market forecasting using ANN techniques have been performed during the past three decades (Distelfeld *et al*., 2007).

Back Propagation (BP) algorithm is a popular model in neural network. It's one of most NN algorithms in financial time series forecasting (Zhang and Kandel, 1998). Actually, it's a systematic method for training or learning algorithm of multilayer ANN rather than the network itself (Khan *et al*., 2011). There are many successful applications for BP neural networks in science, engineering and finance (Siddique and Tokhi, 2001).

The BP uses the gradient descent method for minimization the error function in weights space. We have to guarantee that the error function is differentiable because this method requires computation of the gradient of the error function at each iteration step (Benvenuto and Piazza, 1992). Therefore, we must utilize a kind of activation function. BP algorithm is applicable to all of other kinds of proposed activation functions. Many applications can be formulated by using a BP network and the methodology has been a model for most multilayer neural networks.

Testing-and-acceptance approach: When scientists want to analyse a time series using ICA, they frequently have to choose from the data set a number of such dominant components that will contain the most information regarding the primary aspect of the time series data. Due to those dominant components, the scientist can discover key information about the stochastic mechanism by which the observed series are generated. It is also possible to understand better the workings behind such a series. This is the reason the analysis is made easier due to the studies which focus on such dominant components (Cheung and Xu, 1999). One of the ways to take in order to select the dominant components is by carrying out two separating steps in the procedure below:

Step 1: Make a list of all the ICs in an appropriate order .

Step 2: Choose several components from the beginning of the sequence as the dominant ones (Cheung and Xu, 2001).

Many studies point to the conclusion that the Testing-and-Acceptance (TnA) approach proposed by Cheung and Xu (2001) using Relative Hamming Distance (RHD) reconstruction error are indeed useful tool in ordering of the ICs and in the identification and elimination of the noise IC. The usefulness of the RHD reconstruction error lies in the possibility of accessing the similarity among the initial variables as well as their respective reconstructed variables.

With the RHD value nearing zero, a higher similarity between the original and their corresponding reconstructed variables is demonstrated. This means that more features of the original variables are retained in the corresponding ICs used to reconstruct them and conversely, less of the effective information is preserved in the eliminated IC. However with the RHD value far from zero there is little similarity between the original and corresponding reconstructed variables, i.e., there is more effective information stored in the eliminated IC. Thus the aim is to find the RHD value closest to zero since the noise IC is the corresponding eliminated IC (Liu and Wang, 2011).

Fast ICA-BP forecasting model: The integration of different prediction techniques is increasingly popular in recent studies because each one of these techniques has advantages and disadvantages. ANN technique is regarded as more suitable for stock market forecasting than others. In the modeling of stock market forecasting using NN, one of the key problems is that financial time series contain inherent noise. If the forecasting model is built without taking into account their potential noise content this may weaken the ability to generalize to the test set. Moreover the noise in the data could lead to the over-fitting or under-fitting problem (Cao, 2003). Therefore analysing the raw dataset is essential task during the modeling process. ICA is used for this purpose. Detecting and removing the noise is very important but difficult task when building a forecasting model. TnA technique is used to solve this problem.

To minimise the impact of noise, a three-stage approach by integrating FastICA, TnA and BP techniques, called FastICA-BP is proposed in this work in order to obtain an efficient stock market forecasting model. In this model, FastICA is firstly used to analyse the raw dataset to get components which are independent of each other. Secondly, TnA has been approach applied to identify and remove the IC representing the noise using relative hamming distance reconstruction error. And finally, BP algorithm has been used to predict the stock market using the filtered components.

RESULTS AND DISCUSSION

Empirical research: In financial field the historical data of the last trading days including, daily open price, daily highest price, daily lowest price daily closing price, daily volume and daily turnover are commonly used in stock market forecasting. Actually, indicators including, daily open price, daily highest price, daily lowest price and daily closing price are sufficient to forecast the stock market.

In this study, the previous indicators of the last trading days have been used as an input to the BP algorithm, the closing price of the next trading day as the output to forecast stock market price. To verify the efficiency of the proposed model, the aforementioned steps have been applied to construct three different models of financial time series as follows:

Model I: N original financial time series represent the raw dataset of the stock market observations (without preprocessing) conducted as inputs to BP algorithm.

Model II: N independent components dataset which represents the dataset in Model I after analysing using FastICA algorithm (without removing the noise IC) conducted as inputs to BP algorithm.

Model III: N-1 independent components dataset which represent the dataset of Model II after removing the noise IC conducted as inputs to BP.

Model III represents our proposed model that we seek to assess its performance to forecast the stock markets through practical experiments conducted in this study.

To evaluate the performance of the proposed model, dataset from Al Rajhi Islamic bank has been used as illustrated example. Figure 1 shows four financial time series raw dataset representing daily open price, daily

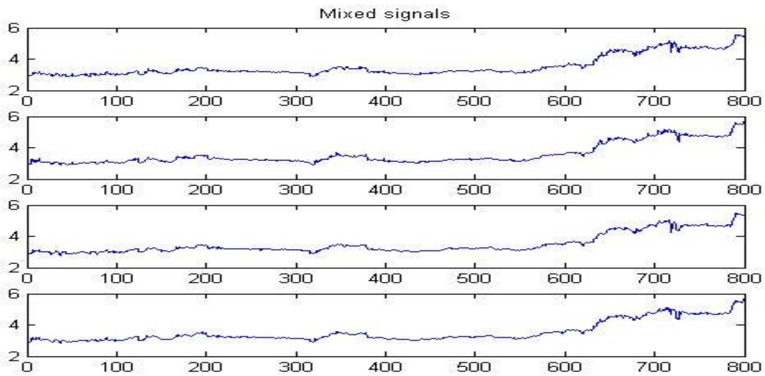

Fig. 1: Original four time series data of Al Rajhi bank from October 11, 2010 to December 31, 2013

highest price, daily lowest price and daily closing price respectively have been collected from October 11, 2010 to December 31, 2013 of Al Rajhi bank trading day, each of size 800 data. Dataset has been divided into two sets, the first 500 data points have been used as the training sample and the remaining 300 data points have been used as the testing sample.

The three chosen models have been applied to the same selected dataset to forecast the closing price of the next trading day. To verify the efficiency of each model, the following statistical criteria have been used: Root Mean Square Error (RMSE), Mean Absolute Error (MAE), Normalized Mean Square Error (NMSE), Mean Absolute Percentage Error (MAPE) as well as the Correlation Coefficient (R). The smallest values of RMSE, MAE, NMSE, MAPE and the largest value of R represent the best performance of the forecasting model. The definitions of these criteria can be found in Eq. 4. Own elaboration based on (Agarwal, 2002):

$$\text{RMSE} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(A_i - P_i)^2} \tag{4}$$

$$\text{MAE} = \frac{1}{N}\sum_{i=1}^{N}\left|A_i - P_i\right|$$

$$\text{SE} = \frac{1}{\sigma^2 N}\sum_{i=1}^{N}(A_i - P_i)^2$$

where:

$$\sigma^2 = \frac{1}{N-1}\sum_{i=1}^{N}(P_i - \bar{P})^2$$

$$\text{MAPE} = \left(\frac{1}{N}\sum_{i=1}^{N}\left(\frac{A_i - P_i}{A_i}\right)^2\right)\times 100$$

$$R = \sum_{i=1}^{N}(A_i - \bar{A})(P_i - \bar{P}) \Big/ \sqrt{\sum_{i=1}^{N}(A_i - \bar{A})^2 \sum_{i=1}^{N}(P_i - \bar{P})^2}$$

Table 1: The forecasting results of the actual and predicted closing price of testing dataset in Model I

Criteria	RMSE	NMSE	MAE	MAPE	R
Value	0.609564	0.020550	0.503371	0.007754	0.990962

Table 2: The forecasting results of actual and predicted closing price of testing dataset of Model II

Criteria	RMSE	NMSE	MAE	MAPE	R
Value	0.546690	0.0168251	0.441836	0.006779	0.9916855

Own elaboration

It should be noted that A_i and P_i are the actual and predicted values respectively, that $\bar{A}$ and P stand for the mean of A and P respectively and that N is the total number of data points. The practical experiment procedures have been applied on the three different models as follows:

Model I: The previous four financial time series have been conducted as an input to the BP algorithm in order to forecast the closing price of the next trading day. Table 1 shows the values of erroe between the actual and predicted closing price of the testing dataset during the period time from October 14, 2012 to December 31, 2013.

Figure 2 depicts the comparison between the actual and predicted closing price of the testing dataset in Model I. The figure shows that the predicted closing price is approaching to the actual closing price to some extent. This is due to the efficiency of the BP algorithm in forecasting.

Model II: Original financial time series dataset has been analysed using the FastICA algorithm to obtain four statistically mutually ICs denoted by IC_1, IC_2, IC_3 and, IC_4, respectively (Fig. 3). Table 2 concludes the forecasting results between the actual and predicted closing price of the testing dataset in Model II. It should be noted that the inputs to the BP algorithm are the ICs instead of the raw dataset.

Table 3: The RHD reconstruction error values

The IC Excluded	The IC Included	RHD Value
IC1	IC2 IC3 IC4	1.9973
IC2	IC1 IC3 IC4	1.8689
IC3	IC1 IC2 IC4	0.9585
IC4	IC1 IC2 IC3	2.7996

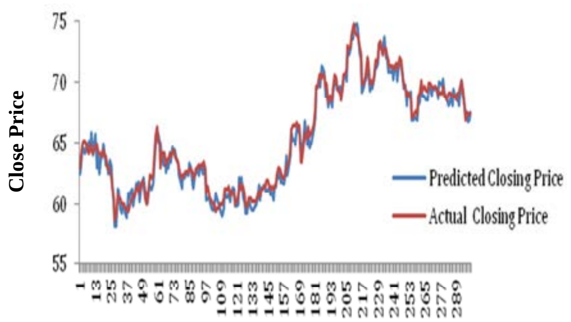

Fig. 2: Comparison between actual and predicted closing price of testing dataset in Model I

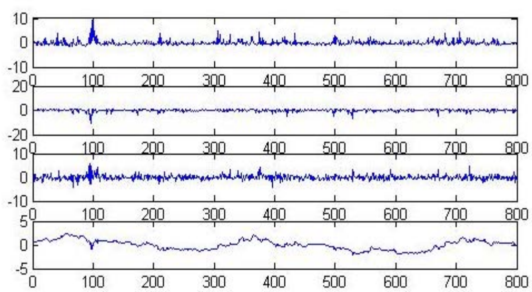

Fig. 3: Four ICs obtained from analysed original raw dataset in Model I using FastICA algorithm

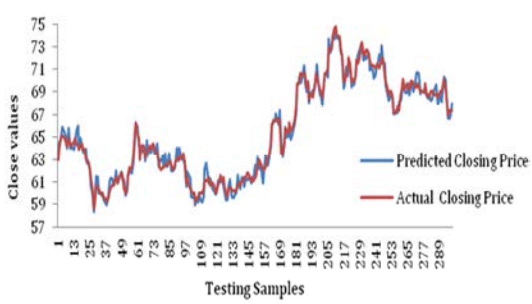

Fig. 4: Comparison between actual and predicted closing price of testing dataset in Model II
Own elaboration

Figure 4 depicts the comparison between the actual and predicted closing price of the testing dataset in Model II. It is clear that the predicted closing price is closer to the actual closing price than it is in the Model I. This is due to use of FastICA algorithm as a preprocessing step to analyse the raw dataset.

Table 4: The forecasting results of actual and predicted closing price of testing dataset of Model III

Criteria	RMSE	NMSE	MAE	MAPE	R
Value	0.347491	0.0067664	0.282554	0.004339	0.996617

Table 5: Total forecasting results of the three different models.

Model Criteria	Model I	Model II	Model III
RMSE	0.609564	0.5466900	0.3474910
NMSE	0.020550	0.0168251	0.0067664
MAE	0.503371	0.4418360	0.2825540
MAPE	0.007754	0.0067790	0.0043390
R	0.990962	0.9916850	0.9966170

Own elaboration

In the next step, the TnA approach has been applied on the four ICs in Model II to identify the noise IC using RHD reconstruction error. Table 3 concludes the RHD values of each step. In this procedure each component is assumed as the last one in the ordering and excluded in reconstructing the mixture matrix. One component has been excluded from each iteration step and the remaining components have been used to reconstruct the original data.

According to Table 3, it is clear that the smallest RHD value was in the third stage when IC3 was eliminated. This means that IC3 contains fewer features of the original data, i.e., IC3 represent the noise IC and therefore it should be removed. The remaining three ICs have been used as the dataset of Model III (Fig. 5).

Model III: The remaining three ICs of Model II dataset after removing the noise IC has been used as an input to the BP algorithm. Table 4 illustrates the forecasting results between the actual and predicted closing price of the testing dataset in Model III. One should be noted that the inputs of the BP algorithm are the three filtered ICs after removing the noise IC.

Figure 6 demonstrates the comparison between the actual and predicted closing price of the dataset in Model III. It shows that the predicted closing price is almost identical with the actual closing price to some extent. This due to the analysis and elimination the noise from the raw dataset by integrating FastICA with the TnA approaches.

To save time, Table 5 shows the summary of forecasting results in previous three different models (Table 3). Total forecasting results of the three different models. It shows that the proposed model (FastICA-BP), namely Model III has fewer errors and more accuracy among the other models. Subsequently, the proposed model has more efficiency and the best performance than the other two models in stock market forecasting.

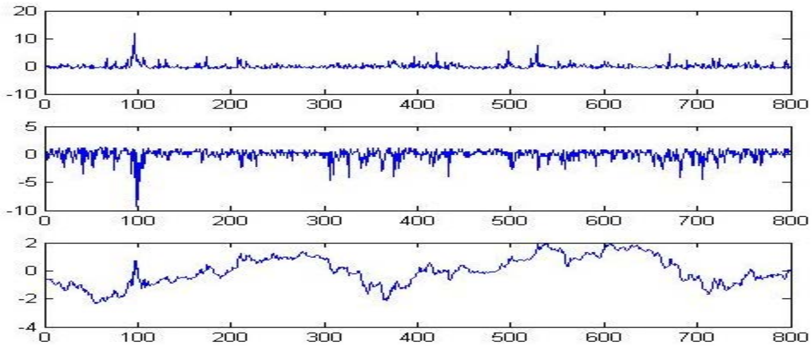

Fig. 5: Rest three ICs of Model II dataset after removing the noise IC.

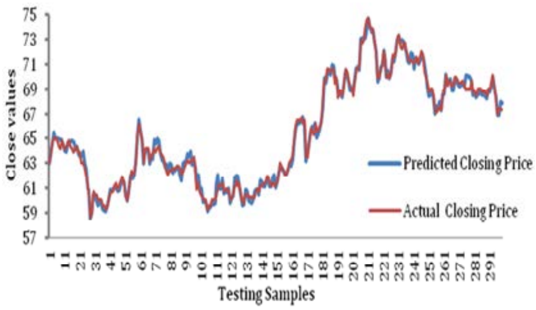

Fig. 6: Comparison between actual and predicted closing price of testing dataset in Model III

CONCLUSION

In this study we present a best model of forecasting the Gulf Cooperation Council (GCC) stock market by combining FastICA, TnA and BP techniques (FastICA-BP) model. In this model, FastICA is firstly used to analyse the raw dataset to get components which are independent of each other. Secondly, TnA approach has been applied to identify and remove the IC representing the noise using RHD reconstruction error. And finally, BP algorithm has been used to predict the stock market using the filtered components. To evaluate the performance of the proposed model, Al Rajhi Islamic bank has been used as illustrative example. The proposed model has been compared with two different models: BP model using original data set (Model I) and BP model using non-filtered dataset (Model II). The experimental results show that the proposed model the (FastICA-BP), namely Model III gives predictive results with less error and more accuracy in comparison with the other models. This means that the proposed model is more efficient and best performing in predicting stock markets from the rest models.

REFERENCES

Agarwal, N., 2002. Neural network based modeling and simulation for the optimization of safety logic. Master Thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts.

Benvenuto, N. and F. Piazza, 1992. On the complex backpropagation algorithm. IEEE. Trans. Signal Process., 40: 967-969.

Bingham, E., 2003. Advances in independent component analysis with applications to data mining. M.sc Thesis, Department of Computer Science and Engineering, Helsinki University of Technology, Espoo, Finland. https://aaltodoc.aalto.fi/handle/123456789/2141

Cao, L., 2003. Support vector machines experts for time series forecasting. Neurocomputing, 51: 321-339.

Cheung, Y.M. and L. Xu, 1999. An empirical method to select dominant independent components in ICA for time series analysis. Proceedings of the International Joint Conference on Neural Networks (IJCNN'99), July 10-16, 1999, IEEE, Hong Kong, ISBN:0-7803-5529-6, pp: 3883-3887.

Cheung, Y.M. and L. Xu, 2001. Independent component ordering in ICA time series analysis. Neurocomputing, 41: 145-152.

Distelfeld, A., I. Cakmak, Z. Peleg, L. Ozturk and A.M. Yazici et al., 2007. Multiple QTL-effects of wheat Gpc-B1 locus on grain protein and micronutrient concentrations. Physiol. Plant., 129: 635-643.

Hyvarinen, A. and E. Oja, 2000. Independent component analysis: Algorithms and applications. Neural Networks, 13: 411-430.

Hyvarinen, A., 1997. A family of fixed-point algorithms for independent component analysis. Proc. IEEE Int. Conf. Acoust. Speech Signal Process., 5: 3917-3920.

Hyvarinen, A., 1999. Fast and robust fixed-point algorithms for independent component analysis. IEEE Trans. Neural Networks, 10: 626-634.

Hyviirinen, A., J. Karhunen and E. Oja, 2001. Independent Component Analsis. John Wiley & Sons, New York, USA.,.

Khan, Z.H., T.S. Alin and M.A. Hussain, 2011. Price prediction of share market using Artificial Neural Network (ANN). Int. J. Comput. Appl., 22: 42-47.

Khaze, S.R., M. Masdari and S. Hojjatkhah, 2013. Application of artificial neural networks in estimating participation in elections. Int. J. Inf. Technol. Model. Comput. (IJITMC.), 1: 23-31.

Liu, H. and J. Wang, 2011. Integrating independent component analysis and principal component analysis with neural network to predict Chinese stock market. Math. Prob. Eng., 2011: 1-15.

Lu, C.J., T.S. Lee and C.C. Chiu, 2009. Financial time series forecasting using independent component analysis and support vector regression. Decis. Support Syst., 47: 115-125.

Siddique, M.N.H. and M.O. Tokhi, 2001. Training neural networks: Backpropagation vs. genetic algorithms. Proceedings of the International Joint Conference on Neural Networks, July 15-19, Washington, DC, USA., pp: 2673-2678.

Zhang, Y.Q. and A. Kandel, 1998. Compensatory Genetic Fuzzy Neural Networks and their Applications. Vol. 6, World Scientific, USA., pp: 18-28.