Forecasting Stock Exchange Movements Using Artificial Neural Network Models and Hybrid Models

Erkam GÜREŞEN and Gülgün KAYAKUTLU

Istanbul Technical University, Department of Industrial Engineering,

Maçka, 34367 Istanbul, Turkey.

erkamguresen@gmail.com

kayakutlu@itu.edu.tr

Abstract: Forecasting stock exchange rates is an important financial problem that is receiving increasing attention. During the last few years, a number of neural network models and hybrid models have been proposed for obtaining accurate prediction results, in an attempt to outperform the traditional linear and nonlinear approaches. This paper evaluates the effectiveness of neural network models; recurrent neural network (RNN), dynamic artificial neural network (DAN2) and the hybrid neural networks which use generalized autoregressive conditional heteroscedasticity (GARCH) and exponential generalized autoregressive conditional heteroscedasticity (EGARCH) to extract new input variables. The comparison for each model is done in two view points: MSE and MAD using real exchange daily rate values of Istanbul Stock Exchange (ISE) index XU10).

1. Introduction

The financial time series models expressed by financial theories have been the basis for forecasting a series of data in the twentieth century. Yet, these theories are not directly applicable to predict the market values which have external impact. The development of multi layer concept allowed ANN (Artificial Neural Networks) to be chosen as a prediction tool besides other methods. Various models have been used by researchers to forecast market value series by using ANN (Artificial Neural Networks). Engle (1982) suggested the ARCH(p) (Autoregressive Conditional Heteroscedasticity) model, Bollerslev (1986) generalized the ARCH model and proposed the GARCH (Generalized ARCH) model. By considering the leverage effect limitation of the GARCH model, the EGARCH (Expo-

Please use the following format when citing this chapter:

Güreşen, E. and Kayakutlu, G., 2008, in IFIP International Federation for Information Processing, Volume 288; *Intelligent Information Processing IV*; Zhongzhi Shi, E. Mercier-Laurent, D. Leake; (Boston: Springer), pp. 129–137.

nential GARCH) model was proposed (Nelson 1991). Despite the popularity and implementation of the ANN models in many complex financial markets directly, shortcomings are observed. The noise that caused by changes in market conditions, it is hard to reflect the market variables directly into the models without any assumptions (Roh 2007). During the last few years research is focused on improving the ANN's prediction performance.

The objective of this study is to compare classical ANN models and new ANN methodologies with hybrid ANN models, such as GARCH-ANN and EGARCH-ANN models. The methods are compared by using MSE (Mean Square Error), MAD (Mean Absolute Deviation) and % MAD (Mean Absolute % Deviation).

The remaining sections of this paper are organized as follows: Section 2 provides a brief review of related studies. Section 3 introduces the models used in this study and Section 4 provides results of each model using daily exchange rates of Istanbul Stock Exchange (ISE) index XU100. Final section concludes the study with future researches.

2. Brief review of research on time series

ANN models have been used by researchers. A brief literature survey is given in Table 1. This survey clearly shows that ANN methods outperform the classical methods. Hybrid methods that use both classical methods with ANN have potential to avoid deficiencies in classical methods.

Many researchers pointed that hybrid methods are promising for future studies and with using hybrid methods advantages of each method can combine.

Table 1. Financial Time Series Researches (ANN and Hybrid Models)

Date	Researchers	Used Method	Data Years	Data Type	Goal	Prediction Period	Results
2007	Preminger and Franck	Robust Liner Autoregressive Robust Neural Nertwork	1971-2004	GBP/$ JPY/$	To optain better results than Standart linear autoregressive and Neural Network	1-3-6 months	Robust models are better than standart models but still are not better than RW (Random Walk)
2007	Hamzaçebi and Bayramoğlu	ARIMA ANN	2002-2006	ISE-XU100	To compare ARIMA and ANN	Daily	ANN has better results
2007	Pekkaya and Hamzaçebi	LR (Linear regression) ANN	1999-2006	YTL/USD	To compare the forecasts using macro economic variables	Monthly	ANN gives better results and predicts two important breaking point with 6.611 % error
2007	Roh	ANN,EWMA (Exponentially Weighted Moving Average) GARCH, EGARCH	930 trading days	KOSPI 200	To compare ANN with hybrid models	Daily	Classical ANN outperforms NN-EWMA NN-EGARCH For periods shorter than a month 100 % direction prediction and for periods shorter than 160 days min 50 % direction prediction, NN-GARCH For periods shurter than a month 100 % direction prediction and for periods shorter than 160 days min 50 % direction prediction
2007	Kumar and Ravi	ANN Fuzzy Logic Cased-Based Resoning Decision Trees Rough Sets			Review- Bankruptcy prediction (128 paper)		SVM outperforms logistic regression and BPNN Rough set based Ap. outferforms logistic regression and decision tree Logistic regression, LDA, QDA, FA clearly outperformed by ANN Hybrid methods combine the advantages and promising for future researches
2007	Celik and Karatepe	ANN	1989-2004	Monthly banking sector data series	Crises prediction		Financial ratios successfully predicted for 4 months
2005	Ghiassi, Saidane and Zimbra	ANN, ARIMA DAN2 (Dynamic Architecture for ANN)		Time series used in literature	To compare the methods		DAN2, is an alternative of ANN and gives better result and only needs to choose the inputs
2006	Menezes and Nikolaev	Genetic Programming (GP) Polynomial Genetic Programming (PGP)		Time series used in literature	To compare the methods		The polynomials in time series are found and promising for future researches
2007	Zhang and Wan	Fuzzy Interval NN (FINN)	1998-2001	JPY/USD GBP/USD	Exchange prediction	6 weeks	Promising for future researches
2007	Hassan, Nath and Kirley	Hidden Markov Model (HMM), ANN Genetik Algorithm (GA)	2003-2004	Stocks; Apple Computer Inc., IBM, Dell Inc.	Exchange prediction	5 weeks	Hybrid model is better than HMM and ARIMA
2005	Yümlü, Gürgen and Okay	Mixture of Experts (MoE) MLP RNN EGARCH	1990-2002	ISE XU100 daily values	Exchange prediction & To compare the methods	4 years	MoE outperforms the aother models EGARCH is outperformed by all other methods

3. ANN and Hybrid ANN Models

3.1. Multilayer Perceptron (MLP)

This model uses last four values of a time series as inputs and generated by using NeuroSolutions 5.06 software. MLP has two layers using tanh neurons. The number of neurons in each layer and learning rate are calculated by genetic algorithm using the same software.

3.2. Lagged Time Series (LTS)

This model is generated by using NeuroSolutions 5.06 software to use lagged values of the financial time series. LTS has 2 layers with tanh neurons and each layer have lagged connections. The number of neurons in each layer and learning rate are calculated by genetic algorithm using the same software.

3.3. Recurrent Neural Network (RNN)

This model is generated by using NeuroSolutions 5.06 software to have 2 layers with tanh neurons and each layer consisting of recurrent connections. The number of neurons in each layer and learning rate are calculated by genetic algorithm using the same software.

3.4. Dynamic Architecture for Artificial Neural Networks (DAN2)

This model is developed by Ghiassi and Saidane (Ghiassi and Saidane 2005) and compared with the classical ANN models using a known time series (Ghiassi et al. 2005). Figure 1 shows the structure of DAN2.

DAN2 uses all input data at a time to train the network. Training begins with a special F_0 node captures the linearity using classical linear regression. The training process stops when a desired level of accuracy is reached. Each time a nonlinear relation is hit, a new hidden layer is added. Each hidden layer has 4 nodes: one C node, one CAKE node (in Figure 1, F nodes) and two CURNOLE nodes (in Figure 1, G and H nodes). A CAKE (Current Accumulated Knowledge Element) node captures the previous layers using the CAKE node at the previous layer.

With a linear combination of CURNOLE (Current Residual Nonlinear Element) nodes, C node and previous CAKE node, existing CAKE node provides the results. Until the desired level of accuracy reached new hidden layers continue to be added to the model. After the special linear layer (F_0) DAN2 uses α_i's where α_i is the angle between the observation vector *i* and a defined reference vector. DAN2 uses the trigonometric transfer functions to capture the nonlinearity. Each *G* and *H* nodes at layer *k* uses the given formula:

$$G_k(X_i) = \text{Cosine}(\mu_i \times \alpha_i)\ ,\ H_k(X_i) = \text{Sine}(\mu_i \times \alpha_i) \qquad [1]$$

Using the given formula of $G_k(X_i)$ and $H_k(X_i)$ we can use the following formula for F nodes:

$$F_k(X_i) = a_k + b_k\, F_{k-1}(X_i) + c_k \text{Cosine}(\mu_i \times \alpha_i) + d_k \text{Sinse}(\mu_i \times \alpha_i) \qquad [2]$$

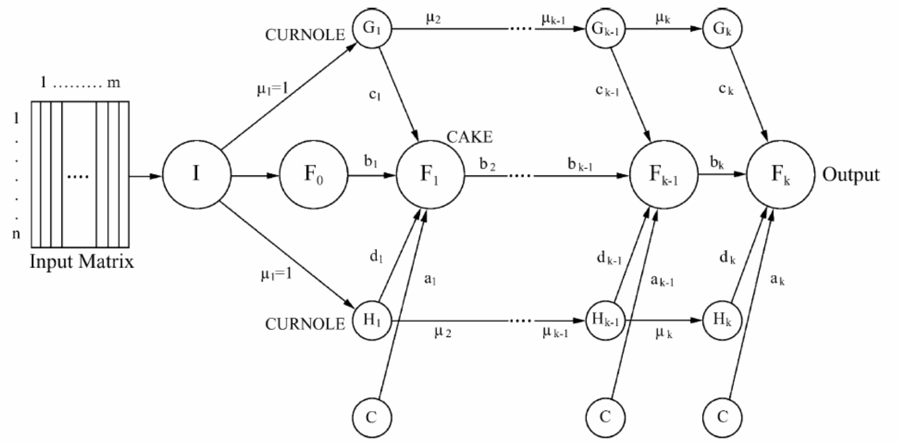

Fig. 1. The DAN2 Network Architecture (Ghiassi and Saidane 2005)

3.5. GARCH - ANN Models

Most of the financial series models are known to be easily modelled by GARCH(1,1), so this research uses the extracted variables from GARCH(1,1) as Roh suggests (Roh 2007). The GARCH(1,1) has the following formula:

$$\sigma_t^2 = \alpha_0 + \alpha_1 \varepsilon_{t-1}^2 + \beta_1 \sigma_{t-1}^2 \qquad [3]$$

Where σ_t is volatility at t, α_0 is the non-conditional volatility coefficient, ε_{t-1}^2 residual at $t-1$, σ_{t-1}^2 is the variance at $t-1$.

The newly extracted variables are as follows (Roh 2007):

– $\sigma_t^{2'} = \beta_1 \sigma_{t-1}^2$
– $\varepsilon_{t-1}^{2'} = \alpha_1 \varepsilon_{t-1}^2$

We use these new variables as additional inputs for every type of ANN given above.

3.6. EGARCH - ANN Models

EGARCH has the leverage effect with the following formula:

$$\ln \sigma_t^2 = \alpha + \beta \ln \sigma_{t-1}^2 + \gamma \left(\left| \frac{\varepsilon_{t-1}}{\sigma_{t-1}} - \sqrt{\frac{2}{\pi}} \right| \right) + \omega \left(\frac{\varepsilon_{t-1}}{\sigma_{t-1}} \right) \qquad [4]$$

Where α is the non-conditional variance coefficient, $\ln \sigma_t^2$ is the log value of variance at $t-1$, $(| \varepsilon_{t-1}/\sigma_{t-1} - \sqrt{(2/\pi)} |)$ is the asymmetric shock by leverage effect, and $(\varepsilon_{t-1}/\sigma_{t-1})$ is the leverage effect. The newly extracted variables are as follows (Roh 2007):

– $\ln \sigma_t^{2'} = \beta \ln \sigma_{t-1}^2$
– LE (leverage effect) $= \gamma(| \varepsilon_{t-1}/\sigma_{t-1} - \sqrt{(2/\pi)} |)$
– L(leverage) $= \omega(\varepsilon_{t-1}/\sigma_{t-1})$

4. Forecasting ISE XU100 Index

In this research daily stock exchange rates of ISE index XU100 from January 2003 to March 2008 are used. Graph of the data is given in figure 2. First 1132 days are used for training and cross validation and last 160 used for testing. For hybrid models also new variables extracted from GARCH and EGARCH are calculated using MS Excel. For MLP, LTS, RNN, GARCH-MLP, GARCH-LTS, GARCH-RNN, EGARCH-MLP, EGARCH-LTS and EGARCH-RNN NeuroSolutions 5.06 software is used. For calculating DAN2, GARCH-DAN2 and EGARCH-DAN2 MS Excel is used. Results are given in table 2. GARCH-DAN2 have the smallest training MSE and MAD, followed by EGARCH-DAN2 and DAN2. In all the other hybrid models, training MSE and MAD values are increased. However, GARCH-DAN2 and EGARCH-DAN2 have smaller training MSE and MAD, DAN2 has smaller testing MSE and MAD. DAN2 based neural networks outperformed the other neural networks. Hybrid RNN models decrease the training error but increase the testing errors.

Fig. 2. ISE XU100 closing values from January 2003 to March 2008

Table 2. Results of ANN and Hybrid Models

	Training			Test		
	MSE	**MAD**	**MAD %**	**MSE**	**MAD**	**MAD %**
MLP	332,121.4	431.074	2.02378	5,540,545.9	2,042.031	3.87061
LTS	4,040,290.2	1,270.136	8.091805	4,053,666.2	3,114.716	6.021672
RNN	2,215,589.2	1,073.526	6.182388	30,728,867.0	4,748.948	8.847816
DAN2	262,130.4	370.661	1.408297	1,176,015.7	840.700	1.679289
GARCH-MLP	468,823.2	514.225	2.627341	7,124,780.4	2,317.443	4.38835
EGARCH-MLP	450,787.2	512.206	2.705861	8,651,756.5	2,547.234	4.797743
GARCH-LTS	4,793,112.9	1,344.516	7.021188	82,679,183.4	8,259.680	15.56614
EGARCH-LTS	7,268,783.3	1,771.432	9.802969	86,388,074.0	8,383.227	15.77058
GARCH-RNN	1,588,036.6	839.538	4.413098	40,952,240.9	5,621.619	10.52457
EGARCH-RNN	2,331,406.0	806.284	4.545228	46,952,272.1	5,970.485	11.15228
GARCH-DAN2	261,378.6	370.218	1.4039	1,178,820.5	842.373	1.682031
EGARCH-DAN2	261,918.2	370.416	1.405955	1,177,072.3	841.188	1.680164

5. Conclusion

This study is in search for reducing the shortcomings of using ANN in predicting the market values. With this aim Hybrid models are developed and investigated. In order to present the differences in accuracy of prediction, all the models are applied on the same set of data retrieved from Istanbul Stock exchange.

This study shows that DAN2 is powerful neural network architecture. Hybrid models using GARCH and EGARCH can decrease the training error but do not guarantee a decrease in testing errors. The lowest error is achieved by DAN2 based hybrid model, which also shows that DAN2 has greater noise tolerance.

The achieved results indicate that DAN2 model is to be focused in the future studies to improve the noise tolerance. More attention is to be given to the hybrid models in defining the hybridization procedure clearly.

References

1. Bollerslev, T., "Generalized autoregressive conditional heteroscedasticity", Journal of Econometrics, vol 31, 1986, p. 307-327.
2. Celik A. E., Karatepe Y., "Evaluating and forecasting banking crises through neural network models: An application for Turkish banking sector", Expert systems with Applications, vol. 33, 2007, p. 809-815.
3. Engle R. F., "Autoregressive conditional heteroscedasticity with estimator of variance of United Kingdom inflation", Econometrica, 1982, 50(4), p. 987-1008.
4. Ghiassi M., Saidane H., "A dynamic architecture for artificial neural networks", Neurocomputing, vol. 63, 2005, p. 397-413.
5. Ghiassi M., Saidane H., Zimbra D.K., "A dynamic artificial neural network model for forecasting time series events", International Journal of Forecasting, vol. 21, 2005, p. 341-362.
6. Hassan M. R., Nath B., Kirley M., "A fusion model of HMM,ANN and GA for stock market forecasting", Expert Systems with Applications, vol. 33, 2007, p. 171-180.
7. Kumar P. R., Ravi V., "Bankruptcy prediction in banks and firms via statistical and intelligent techniques – A rewiev", European Journal of Operational Research, vol. 180, 2007, p. 1-28.
8. Menezes L. M., "Forecasting with genetically programmed polynomial neural networks", International Journal of Forecasting, vol. 22, 2006, p. 249-265.

Nelson D. B., "Conditional heterosdasticity in asset returns: a new approach", Econometrica, 1991, 59(2), p. 347-370.

9. Pekkaya M., Hamzaçebi C., "Yapay sinir ağları ile döviz kuru tahmini üzerine bir uygulama", YA/EM 27. National Congress, Izmir 2007, p. 973-978.
10. Preminger A., Raphael F., "Forecasting exchange rates: A robust regression approach", International Journal of Forecasting, vol. 23, 2007, p. 71-84.
11. Roh T. H., "Forecasting the volatility of stock price index", Expert Systems with Applications, vol. 33, 2007, p. 916-922.
12. Zhang Y., Wan X., "Statistical fuzzy interval neural networks for currency exchange rate time series prediction", Applied Soft Computing, vol. 7, 2007, p.1149-1156.