TURKISH STOCK MARKET ANALYSIS USING MIXTURE OF EXPERTS

M.Serdar Yumlu
Bogazici University
Computer Engineering Dept.
34342 Bebek Istanbul/Turkey
yumlu2@yahoo.com

Fikret S. Gurgen
Bogazici University
Computer Engineering Dept.
34342 Bebek Istanbul/Turkey
gurgen@boun.edu.tr

Nesrin Okay
Bogazici University
Management Department
34342 Bebek Istanbul/Turkey
okay@boun.edu.tr

ABSTRACT

Istanbul Stock Exchange (ISE) stock market is not an efficient market. In this paper, we show how localized and global artificial neural network (ANN) models are used for risk estimation of asset returns. Mixture of Experts (MoE) and Recurrent Neural Networks (RNN) are more powerful to say that Efficient Market Hypothesis (EMH) is violated. ISE index XU100 is studied using daily data over a 14-year period using MoE and RNN neural networks and also Glosten-Jaganathan-Runkle (GJR**)** volatility models. The results suggest that localized neural approaches have the strength in modeling the risk in stock market time series data set of XU100.

INTRODUCTION

Time series prediction is fitting a model to the data that is not constant in time. Forecasting is finding an approximate mapping function between the input and the output data. Financial Time Series Forecasting aims to find underlying patterns, trends, cycles and forecast future using historical and currently observable data. ANNs are one of the most innovative approaches to prediction problems and in this study we propose global and local neural approaches to capture the stock market time series data and its predictive content.

Predicting time series preserves conditional dependence; therefore we use a financial database dependent on historical and current observations. In the market, not all of the information is publicly available. *Efficient Market Hypothesis* (EMH) states that all methods of deciding when to buy and sell, over a particular time period and using past prices to evaluate the decisions, are inferior to the strategy of buying at the beginning of the period and selling at its conclusion (Fama, 1970). The alternative approach, widely accepted by the traders' environment, is the belief that the stock market is predictable in the sense of technical and fundamental analysis methods. We are against the EMH approach and support this idea using ANNs for local and global modeling.

As Chan presented Sharpe Ratio maximization can be used as a performance measure and Fourier based recurrent network architectures may have signal generation power [1]. Hellstrom in his studies introduced new concepts about data selection procedure related to the field of financial time series prediction and new performance measures such as Theil coefficient etc [2]. Profit based network training and guidelines for time series prediction has been introduced by Yao [3].

The remainder of this paper proceeds as follows: in the second section, volatility forecasting has been introduced. In the third section we have described our data and the model architecture proposed to capture the data as well as the statistical characteristics of the stock market series. The fourth section is related to the neural modeling techniques that we have used in this study. The fifth section describes the results according to the performance and benchmark measures that we have used in this study to compare with other models. Final section concludes the study with future work.

FINANCIAL TIME SERIES AND VOLATILITY

Stock prices vary with changes in volatilities of the underlying risk factor and as a consequence, accurate prediction of future stock prices requires a forecast of asset return's volatility.

Financial time series exhibit time dependent heteroskedastic variance known as conditional variance (volatility) that is not a directly observable feature.

A volatility model [4, 5], proposed in this paper, is used to forecast risk, and return of assets. These forecasts are used in market risk management, portfolio selection, market timing etc. and by other financial decision makers. Therefore, predictability of the volatility is important: a portfolio manager may want to sell a stock or a portfolio before the market becomes too volatile or a trader may want to know the expected volatility to give the right decision while buying or selling a stock. A risk manager has the right to know that his portfolio may likely to decline in near future. None of the players want a volatile market; estimating the volatility of asset returns, which is a basic risk factor component of the stock market, gives valuable information for the future risk in the market and this will make the players to consider the expected high or low volatility in the market.

Glosten, Jagannathan and Runkle proposed GJR model for modeling the asymmetric behavior in time series [6].

$$h_t = w + \beta h_{t-1} + \alpha\varepsilon_{t-1}^2 + \gamma S_{t-1}^{-1}\varepsilon_{t-1}^2, \tag{1}$$

where $S_t^- = 1$, if $\varepsilon_t < 0$, $S_t^- = 0$ otherwise.

GJR model also allows modeling of different impacts of unexpected positive and negative returns.

DATA

Data used in this paper consist of 3650 daily observations of XU100 index of ISE market covering a 14-year period, from 03 January 1989 to 18 September 2003.

In order to use the data effectively in ANN processing, first we have to keep the series in a constant range by applying normalization. The index price series (y_t) is transformed into continuously compounded return series (r_t) by the formula given below to obtain an accepted stationary series. This continuously compounded returns is shown in Figure 1.

$$r_t = \ln(y_t / y_{t-1}) \tag{2}$$

Several statistical characteristics of asset returns have emerged throughout the years and these have been confirmed by numerous studies done by volatility modeling for predicting financial time series [7,8].

Table 1 reports the statistics for XU100 index return series, obtained from ISE database. The results show that XU100 is negatively skewed, and the kurtosis, which is a measure of the thickness of the tails, is very high. Jarque-Bera test statistic rejects the null hypothesis of normal distribution (Bera) [9]. ARCH (5) tests the ARCH effects using Engle's [4] test.

Financial markets in Turkey are chaotic and show non-linearity with very high noise because of the political indeterminate structure and instability of the economy. Any news, any rumor may cause sudden volatile movements making market volatility a very difficult and a complex problem with all hidden but effecting factors [8].

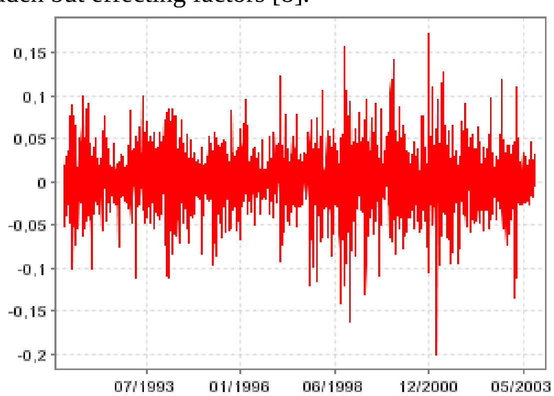

Figure 1 Continuously compounded return series of XU100

Table 1 Descriptive Statistics of daily returns

Observations	3650
Mean	0.0022
Standard Deviation	0.0318
Skewness	-0.1187
Kurtosis	5.4918
Standardized Range	11.657
Q(5)	50.369 (0.1165)
Q(10)	70.075 (0.0043)
$Q^2(5)$	579.262 (0)
$Q^2(10)$	746.101 (0)
ARCH(5)	380.7132 (0)
Jarque-Bera (p-value)	950.5516 (0)
$r_t > 0$ (%)	52.6301
$r_t < 0$ (%)	47.0411
$r_t = 0$ (%)	0.3288

Q (5), and Q2 (5) are the Ljung-Box statistics of the autocorrelation of residuals and squared residuals respectively; J-B is the Jarque-Bera statistic for normality test.

RECURRENT NEURAL NETWORKS (RNN)

RNN allows temporal connections for modeling the relationship in time. Unlike other feed-forward neural network models, RNN has feedback connections that prepare a valuable basis for modeling time series. We have studied Elman RNN networks using Real-time Recurrent Learning (RTRL).

Time can be introduced in the neural architecture in different ways. We may leave the time outside of the neural model as we do in MoE. We have defined a sliding window of last n elements and then use this window to predict the following observation in the time series. We can encode the time as numerical values and use these values in the neural network or as we do in RNN, we can introduce the time into the model as an index of the state of the network. The storage of the states brings the memory and its form. Buffering and weighed buffering techniques can be applied for the form of the memory. We have used RNN for modeling the risk in index return series and use this for prediction purposes.

MIXTURE OF EXPERTS

A mixture of Experts (MoE) model is a divide and conquers approach. In problems, where fitting a global model is inapplicable, dividing the space into regions and fit separate models to each region is a better approach of modeling. MoE proposed by Jacobs [10], is one of the means of combining multiple learners. It seems like voting, but the mechanism, which gives votes to each expert, is not stable over all patterns in MoE and it regards the input patterns while learning.

In MoE modeling, first the data is clustered into groups using Expectation Maximization algorithm. The number of clusters is equal to the number of hidden units in the neural network architecture. Like all divide and conquer approaches, assignment problem also raises in MoE architecture. A gating network is responsible for assigning local experts to pre-defined clusters (regions). Local experts which work in parallel using the same input pattern makes up the first module and the gating network is the second module which produces the outputs using all the outputs of local experts by giving the weights while regarding the input pattern as the local experts do. All experts are trained in parallel on all patterns according to the mean square error measure. The architecture of MoE is given in Figure 2.

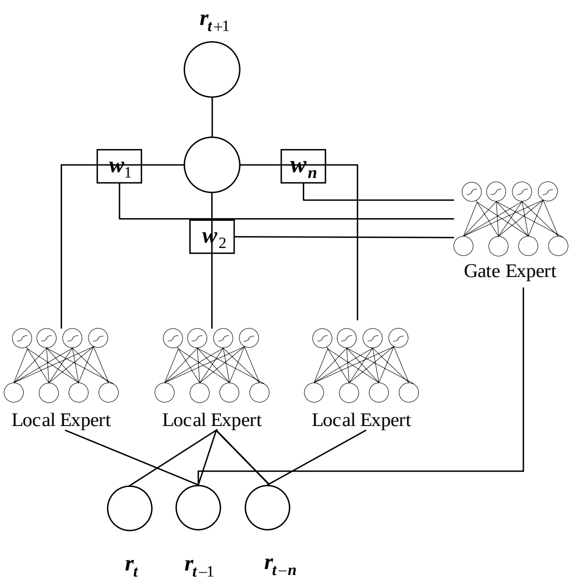

Figure 2 MoE Architecture

RESULTS

A neural network's performance depends on the quality and relevance of the data. We have selected ISE index XU100 for a 14-year period. We have applied a sliding window from the time series as an input to the ANNs and then try to forecast the following derivations in the time series. The data consists of 3650 daily observations and these are divided into three datasets in order to achieve the best generalization. We have trained our models in the train data set, modify and calibrate the parameters of ANNs in the validation data set and then measure the out-of-sample performance of the models in the test (production) data sets. The train, validation and test data sets consist of 2300, 985 and 365 daily observations respectively. Also for ANNs learning rate, momentum rate and number of hidden units that have used in the ANN are important parameters of neural learning. Learning is based on the definition of a suitable error function. Mean square error (MSE) is used for learning and also for comparing the out-of-sample performance of the models we have defined MSE and correlation as regression metrics and also Hit Rate (H_R) for measuring the number of accurate predictions of the sign of the returns and Theil Inequality Coefficient (TIC) for providing a measure of the model performance relative to the naïve predictor. Correlation (CORR) is expected to be as close as to 1 and MSE and TIC should be as close as zero.

Mostly true volatility is accepted as the square of returns (r_t^2). This is not true in real trading environment. Volatility is not such a simple and directly observable measure and only squared returns cannot represent the volatility. But in this study, we will accept the true volatility as RiskMetrics that is proposed by J.P.Morgan. This is a more realistic measure of volatility and will help the neural models to handle the model in a more accurate way.

We have studied one global and one local neural modeling and also GJR based modeling of volatility using XU100 time series data. First we have studied with a 14-year period of data and then studied with several subsets of this data for the analysis of the short-term success of the models that we have used. In Table 2, the performance metrics defined have been reported. The results have shown that localized neural modeling technique MoE has superior characteristics in modeling the stock market time series of XU100 in both the number of the accurate sign of the returns, in fitting the data from the perspective of regression statistics and also in Theil inequality coefficient. RNN also outperformed GJR model when we have used the whole time series data.

Table 2 Out-of-sample performance statistics using data between 1989 and 2003.

	MoE	RNN	GJR
MSE	0.2277	0.3044	0.499
H_R (%)	89.92	81.02	86.4
CORR	0.8158	0.6716	0.648
TIC	0.3882	0.3114	0.511

When we have used a small subset of XU100 data, we have seen that localized neural modeling approach, MoE, has outperformed both RNN and GJR in a more obvious way. We have used data between 03 January 2000 and 18 September 2003. Hit rate performance of 92 percent and also correlation and mean square error measures of perfect fit shows that MoE is more powerful in short-term in Turkish stock market which is highly acceptable in todays economy of Turkey. Market fluctuations do not last more than six months. In short-term local models have the strength in prediction rather than global models.

Table 3 Out-of-sample performance statistics using data between 2000 and 2003.

	MoE	RNN	GJR
MSE	0.02425	0.2082	0.497
H_R (%)	92.5	75.94	87.9
CORR	0.97375	0.5568	0.897
TIC	0.098	0.466	0.447

CONCLUSION AND FUTURE WORK

In this study, we have demonstrated that artificial neural networks are more promising than the GJR model. In stock market time series prediction using XU100 localized modeling approach of MoE is more powerful than global models especially for short-term modeling the success of the localized models have outperformed both RNN and GJR models. ANN is much powerful to say that EMH is violated in XU100. Value-at-Risk (VaR) is very important for portfolio manager in risk management. The uses of volatility in VaR may be incorporated in financial time series prediction problems. Also high frequency data for stock time series, intraday data time series, can be used for very-short-term volatility modeling of stock time series.

ACKNOWLEDGMENTS

This study has been supported by BU03A103 Research Fund.

REFERENCES

1. Chan L., GAO X.: An Algorithm for Trading and Portfolio Management Using Q-learning and Sharpe Ratio Maximization, in proceedings of International Conference on Neural Information Processing, ICONIP 2000, Taejon, Korea, 832-837
2. Hellstrom T., Holmstrom K.: Predictable Pattern in Stock Returns, Technical Report Series, IMA-TOM (1997-09)
3. Yao J.T., Tan C.L.: Guidelines for Financial Forecasting with Neural Networks (2001)
4. Engle R.: Autoregressive Conditional Heteroskedasticity with estimates of the variance of United Kingdom infation, Econometrica. 50 (1982) 987-1007
5. Engle R.F., Ng V.K.: Measuring and Testing the Impact of News on Volatility, The Journal of Finance, Vol. 48, Issue 5 (Dec., 1993) 1749-1778
6. Glosten L, R. Jagannathan, and D. Runkle, On the relation between the Expected Value and the volatility of Nominal Excess Return on Stocks, Journal of Finance, Vol. 48, 1993, 1779-1800.
7. Okay, N., Evidence For Conditional Heteroskedasticity in Turkish Stock Returns, Readings Book of the 1998 ABA National Conference. (1998) 241-247
8. Okay,N. Dynamic predictability of the Istanbul Stock Exchange Market using a Vector Autoregressive Model. Bogazici Journal, 12-31-56
9. Bera A.K.,Higgins M.L., Arch Models. Properties, Estimation and testing. Journal of Economic Surveys. 7 (1993) 305-362
10. Jacobs R.A., Jordan M.I., Nowlan S.J., Hinton G.E.: Adaptive mixture of local experts, Neural Computation, Vol. 3. (1991) 79-87