Professor Melike BİLDİRİCİ, PhD
E-mail:bildiri@yildiz.edu.tr
Yıldız Technical University, Department of Economics
Assoc. Professor Özgür Ömer ERSİN, PhD
E-mail: ozgurersin@beykent.edu.tr
Beykent University, Department of Economics

ASYMMETRIC POWER AND FRACTIONALLY INTEGRATED SUPPORT VECTOR AND NEURAL NETWORK GARCH MODELS WITH AN APPLICATION TO FORECASTING FINANCIAL RETURNS IN ISE100 STOCK INDEX

Abstract. *The study aims to augment commonly applied volatility models with support vector machines and neural networks. Further, fractional integration and asymmetric powers will be introduced. The proposed modeling strategy benefits from neural network based GARCH models and SVR-GARCH models. Following these approaches, the study proposed fractional integration and asymmetric power GARCH structures to obtain SVR-FIAPGARCH and NN-FIAPGARCH models to be evaluated in terms of learning algorithms. Models are evaluated for in-sample and out-of-sample forecasting of daily returns in Istanbul ISE100 stock index. Results suggest several findings: i. fractional integration and asymmetric power structures could be modeled with learning algorithms. ii. volatility clustering, asymmetry and nonlinearity characteristics are modeled more effectively with SVR-GARCH and MLP-GARCH models compared to the GARCH models. iii. SVR-GARCH models provided the lowest error criteria levels in out-of-sample and are closely followed by the MLP-GARCH models.*

Key Words: *Volatility, Stock Returns, ARCH, Fractional Integration, MLP*.

JEL Classification: G12, C32, C52, C53

1 Introduction

Financial time series are subject to volatility clustering and asymmetry which result from different dynamics caused by negative and positive shocks. In financial econometrics, to control volatility, GARCH models are commonly applied in modelling time series. However, GARCH family are criticized in the literature due to their insufficiency in terms of their forecasting capabilities.

The study aims to investigate GARCH models and to augment them with learning algorithms based on neural networks and support vector machines. Donaldson and Kamstra (1997) proposed utilization of neural network models in terms of GJR, GARCH and EGARCH based baseline specifications. Bildirici and Ersin (2009) extended to a family of neural network GARCH models to be estimated with gradient-descent and back propagation learning algorithm cooperation to obtain forecast accuracy improvements. Ou and Wang (2010) and Perez-Cruz et. al. (2003) focused on SVR-GARCH models that benefit from support vector algorithms. The study aims to discuss modeling and estimation of MLP-GARCH and SVR-GARCH models with learning algorithms and to introduce fractional integration and asymmetric power versions of these models to capture different volatility dynamics which gain special importance for financial time series.

2 Literature Review

Econometric modeling of volatility in financial market returns following the ARCH specification of conditional volatility of Engle (1982) and further extended to Generalized ARCH (GARCH) model in Bollerslev (1986) has found many significant applications in light of modeling the distributional aspects such as volatility clustering, heavy tails, non-normal distribution. The Asymmetric GARCH model (AGARCH) aims modeling asymmetric effects of negative and positive shocks; whereas, negative and positive news have different effects on volatility. Accordingly, the Exponential GARCH (EGARCH) model and the GJR-GARCH model are among the main modeling techniques followed in applied econometrics literature. The Asymmetric Power GARCH (APGARCH) model is based on different power transformations without simple squared shocks and conditional variances as in the traditional GARCH models.

Following the achievements in modeling the asymmetry in the volatility, ARCH (GARCH) family models are extended to different nonlinear modeling structures; specifically, regime switching (Hamilton and Susmel, 1994), threshold based regression space division with smooth sigmoid type continuous functions (Hagerud, 1997).

SVM models have gathered a growing interest focusing on various nonlinear regression, time series and forecasting tasks. Kim (2003) utilized SVM's for financial time series forecasting. Cao and Gu (2002) used SVM models with non-stationary time series. Müller et.al. (1997) SVM models for time series prediction for predicting time series with noisy data generating processes and chaotic time series under different loss functions and shown that SVM models provided forecast improvement compared to Radial Basis Functions.

Various attempts had been made to augment the out-of-sample forecasting capabilities of GARCH models since these models had been criticized of performing insufficient in terms of forecasting. The ANN-GARCH (Artificial Neural Network ARCH) developed by Donaldson and Kamstra (1997) augments the GJR model with multi-layer perceptron based neural network architecture with logistic squashing functions to model nonlinearity by benefiting the universal approximation property (Cybenko, 1989) of ANN models. Bildirici and Ersin (2009) augments NN-GARCH models further with various learning algorithms and proposes hybrid models of 9 different GARCH family models to obtain forecast accuracy improvement in stock markets. Ou and Wang (2010) follows a similar hybridization approach by combining the benefits of least squares SVM and GARCH, EGARCH and GJR models to forecast financial volatilities of stock markets. Bildirici and Ersin (2013) and Bildirici and Ersin (2014) integrated STAR and MS type nonlinear models with neural networks where both the conditional mean and variance follows MLP processes.

The literature with respect to SVR models and their applications in Turkish financial markets in terms of GARCH modeling is rather limited. Ince and Trafalis (2006) utilizes SVR and MLP models in which input selection processes are based on ARIMA and VAR modeling techniques. They combine co-integration methodology. An application to Turkish exchange rates is provided. Their findings suggest that SVR outperforms the ANN for two input selection methods. Özdemir et.al. (2011) applied SVM models to forecast the direction of ISE100 stock index and showed that SVM model 76-85% correct classifications. Similarly, Kara et.al. (2011) used ANN and SVM models and found that ANN models provided 75.7% correct classifications; whereas, the correct classifications of SVM was 71.2% for ISE100 daily returns. Our study is different than these studies above since they focus on ANN and SVR implementations of financial variables; whereas, our study focuses on volatility models by evaluating and proposing learning algorithm augmented estimation of SVR-GARCH and MLP-GARCH models with fractional integration and asymmetric power GARCH structures.

In Part 3, MLP and SVR models are evaluated and their relevant hybrid GARCH versions, namely, MLP-GARCH and SVR-GARCH models are discussed. In Part 4, the estimation results and forecast results are evaluated. Models are compared for their in-sample performances and are compared in terms of predictive accuracy for out-of-sample forecasts. Conclusions are given in the last section.

3 Research Methodology and Econometric Models

Time series models may be subject to follow nonlinear processes in different proportions, in the conditional mean and/or in the conditional variance. Accordingly, models investigated in the study are divided into groups by possessing nonlinearity in the conditional mean, variance, or none (or both) in the conditional variance and mean. In the study, the first group of models are the baseline models. The group constitutes the GARCH model, fractionally integrated FI-GARCH, Asymmetric Power APGARCH (Ding, Granger and Engle; 1993) and the fractionally integrated FIAPGARCH models (Baillie, Bollerslev and Mikkelsen; 1996). These models are taken as the baseline family of models. Second group of models utilizes Support Vector Regression type nonlinearity in the conditional mean and the conditional variance processes. The obtained models are SVR-GARCH, SVR-FIGARCH, SVR-APGARCH and SVR-FIAPGARCH.

3.1 MLP and GARCH Models

Multi-Layer Perceptron (MLP), an important class of neural networks consists of a set of sensory units defined with an input layer, one or more hidden layers and an output layer with estimation algorithms that include back-propagation and gradient descent type algorithms (Rumelhart et al., 1986). Artificial Neural Network (ANN) models have significant applications in modeling economic variables and time series. The NN-GARCH and SVR-GARCH models evaluated in the study assume that the conditional mean processes follow a nonlinear process modelled with SVR and ANN processes while the conditional variance processes of a time series is allowed to follow a GARCH process additionally augmented with asymmetric power terms and/or fractionally integrated modeling. Following Donaldson and Kamstra (1997) and Bildirici and Ersin (2009), NN-GARCH models will be derived and further, the fractional integration and asymmetric power terms will be introduced to obtain the NN-FIGARCH, NN-APGARCH and NN-FIAPGARCH variants.

i. NN-GARCH Model

NN-GARCH *(p,q,m)* model is a GARCH*(p,q)* process augmented with single hidden layer neural network as follows,

$$\sigma_t^2 = \alpha_0 + \sum_{i=1}^{p} \alpha_i \varepsilon_{t-i}^2 + \sum_{i=1}^{q} \beta_j \sigma_{t-j} + \sum_{h=1}^{m} \xi_h \psi\left(z_t \lambda_h\right) \tag{1}$$

with the following sigmoid type neuron function,

$$\psi\left(z_t \lambda_h\right) = \left[1 + \exp\left(\lambda_{h,d,w} + \sum_{d=1}^{1}\left[\sum_{w=1}^{m} \lambda_{h,d,w} z_{t-d}^{w}\right]\right)\right]^{-1} \tag{2}$$

$$z_{t-d} = \left[\varepsilon_{t-d} - E(\varepsilon)\right] \Big/ \sqrt{E\left(\varepsilon^{2}\right)} \tag{3}$$

$$\left(1/2\right)\lambda_{h,d,w} \sim uniform\ \left[-1,+1\right] \tag{4}$$

Note that, $\psi(z_t\lambda_h)$ is the sigmoid type activation function of the form 1/(1+exp(-x)) and satisfies the universal approximation conditions derived by Cybenko (1989). $\xi = w$ is the weight vector; defined as $z_t\lambda_h = x_i$ for the input variables. Hence, the parameters in the activation function has the λ_h connection weights as given in equation (4). For a throughout analysis, see Bildirici and Ersin (2009).

ii. NN-APGARCH Model

Asymmetric power GARCH (APGARCH) structure of Ding et.al. (1993) has interesting features in volatility modeling. The NN-APGARCH model belongs to the NN-GARCH models discussed in Bildirici and Ersin (2009) and is an extension of Donaldson and Kamstra (1997) NN-GARCH models. The NN-APGARCH model is obtained by augmenting APGARCH model with artificial neural network architecture and modeling techniques,

$$\sigma_t^{\delta} = \alpha + \sum_{k=1}^{r} \phi_k \left(\left|\varepsilon_{t-k}\right| - \gamma_k \varepsilon_{t-k}\right)^{\delta} + \sum_{t=1}^{q} \beta_i \sigma_{t-i}^{\delta} + \sum_{h=1}^{s} \xi_h \psi(z_t\lambda_h) \tag{5}$$

$$\psi(z_t\lambda_h) = \left[1 + \exp\left(\lambda_{h,d,w} + \sum_{d=1}^{1}\left[\sum_{w=1}^{m} \lambda_{h,d,w} z_{t-d}^{w}\right]\right)\right]^{-1} \tag{6}$$

$$z_{t-d} = \left[\varepsilon_{t-d} - E(\varepsilon)\right] \Big/ \sqrt{E\left(\varepsilon^{2}\right)} \tag{7}$$

$$\frac{1}{2}\lambda_{h,d,w} \sim uniform\ \left[-1,+1\right] \tag{8}$$

where, $\psi(z_t\lambda_h)$ is the logistic function. The NN-APGARCH nests several models. The model reduces to the standard NN-GARCH model for δ =2 and γ_k =0, the NN-NGARCH model for γ_k =0, and the NN-GJR-GARCH model for δ =2 and $0 \le \gamma_k \le 1$; the NN-TGARCH model for for δ =1 and $0 \le \gamma_k \le 1$. For estimation of NN-APGARCH models, Bildirici and Ersin (2009) proposes weight-decay algorithms combined with early-stopping conditions within an algorithm corporation framework, where the back-propagation and conjugate-gradiend-descent algorithms are utilized.

iii. NN-FIAPGARCH Model

In this study, NN-FIAPGARCH model is an augmented version of NN-APGARCH model proposed by Bildirici and Ersin (2009). NN-FIAPGARCH model is also an augmented version of fractionally integrated asymmetric power GARCH model with neural network architecture. The model is defined as,

$$(1-\beta L)\sigma_n^{\delta} = \omega + \left((1-\beta L)-(1-\phi L)(1-L)^{d}\right)\left(|\varepsilon_{n-1}|-\gamma_k\varepsilon_{n-1}\right)^{\delta} + \sum_{h=1}^{s}\xi_{h_t}\psi\left(z_t\lambda_h\right) \quad (9)$$

$$\psi(z_t\lambda_h) = \left[1+\exp\left(\lambda_{h,d,w} + \sum_{d=1}^{1}\left[\sum_{w=1}^{m}\lambda_{h,d,w}z_{t-d}^{w}\right]\right)\right]^{-1} \quad (10)$$

$$z_{t-d} = \left[\varepsilon_{t-d} - E(\varepsilon)\right] \Big/ \sqrt{E(\varepsilon^2)} \quad (11)$$

$$\frac{1}{2}\lambda_{h,d,w} \sim uniform\ [-1,+1] \quad (12)$$

where, $\psi(z_t\lambda_h)$ is the logistic function and h number of neurons. Logistic function belongs to the sigmoid type function family applied in neural network literature. The NN-FIAPGARCH nests several models. The model given in (9)-(12) reduces to the NN-FIGARCH model for restrictions on the power term δ =2 and γ_k =0; the model reduces to NN-FINGARCH model for γ_k =0; and to the NN-FIGJRGARCH model if δ =2 and γ_k is so that it varies between $0 \le \gamma_k \le 1$. Further, the model may be shown as NN-GARCH model if δ =1 in addition to the $0 \le \gamma_k \le 1$ restriction. Furthermore, the model could be represented with short memory characteristics under restrictions on fractional integration parameters. By imposing $d = 0$ to the fractional differentiation parameter the model in Eq. (9) reduces to NN-APGARCH model, the short memory model variant. In this study, only FIGARCH and FIAPGARCH versions will be evaluated.

3.2 Support Vector Machines and GARCH Models

Support Vector Machines (SVM) are learning machines developed by Vapnik (1995) and Cortes and Vapnik (1992). A mainstream analysis of SVM models is given in Christianini and Shaw-Taylor (2000). Similar to backpropagation neural network models with multi-layer architecture (Rumelhart et al. 1986), a Support Vector Machine can be used for pattern recognition and nonlinear regression problems. A SVR or SVM possesses the universal approximation property similar to the MLP model. Similar to the property that a MLP with an optimum number of sigmoidal functions (Cybenko 1989), a SVR with optimum support vectors and basis functions can be shown to provide successful approximations of any borel measurable function (Müller et al. 1997; Vapnik, 1995).

One significant difference between the SVM and MLP is that SVM model uses *structural risk minimization* principle rather than *empirical risk minimization* principle as followed by neural network modeling and regression analysis. In this context, the main idea of a support vector machine is to construct a hyperplane as the decision surface in such a way that the margin of separation between positive and negative observations is maximized (Vapnik, 1995). The induction principle focuses on the error rate in the test sample during the optimization process of a SVM. Further, the error rate is bounded by the sum of the training error rate and a penalty term depending on the Vapnik-Chervonenkis dimension (VC). A SVM is optimum if it produces the value zero for the former (error rate) and minimizes the second term (VC) in the case of separable patterns (Christianini and Shaw-Taylor 2000).

The support vector regression focuses on the construction of the support vector learning algorithm by mapping the original data **x** into a higher dimensional feature space *F* with the use of nonlinear mapping ϕ to obtain a linear regression space.

Given a set of data $G=\left\{\left(x_i,a_i\right)\right\}_{i=1}^{N}$, where x_i is the input vector; a_i is the actual value, and *N* is the total number of data patterns, the SVM regression is stated as,

$$y=f\left(x\right)=w_i\phi_i\left(x\right)+b \tag{13}$$

with $\phi_i:\mathrm{R}^n\rightarrow F, w\in F$. Equation (13) provides a decision surface in the form of a hyperplane. Thus, $\phi_i\left(x\right)$ is the feature of inputs x, b and w_i are the threshold and weights to be estimated by minimizing the regularized risk function,

$$R\left(C\right)=C\frac{1}{N}\sum_{i=1}^{N}L_{\varepsilon}\left(d_i,y_i\right)+\frac{1}{2}\left\|w\right\|^2 \tag{14}$$

and $L_{\varepsilon}\left(d_i,y_i\right)$ defined as,

$$L_{\varepsilon}\left(d_i, y_i\right)=\begin{cases}\left|d_i-y_i\right|-\varepsilon, & \left|d_i-y_i\right|\geq\varepsilon\\ 0 & otherwise\end{cases} \quad (15)$$

C and ε are prescribed parameters, d_i is the actual value and y_i is the estimation value at *i*. In the model, linear regression in a high dimensional feature space corresponds to nonlinear regression in the low dimensional input space. $L_{\varepsilon}\left(d_i, y_i\right)$ in Eq. (15) is the ε-insensitive loss function. The norm of w, $\frac{1}{2}\|w\|^2$ measures the flatness of the function; *C* represents the trade-off between the flatness and empirical risk. Two positive slack variables ξ_i and ξ_i^* represent the distance from actual values to boundary values of the ε-tube. Soft margin computation requires the formulation of Equation (14) in the constrained form,

$$R\left(w,\xi_i,\xi_i^*\right)=\frac{1}{2}\|w\|^2+C\left(\sum_{i=1}^{N}\left(\xi_i-\xi_i^*\right)\right) \quad (16)$$

subject to,

$$\begin{aligned}&w_i\phi_i\left(x\right)+b-d_i\leq\varepsilon+\xi_i^*,\\ &d_i-w_i\phi_i\left(x\right)-b\leq\varepsilon+\xi_i \ ,\\ &\xi_i \ ,\xi_i^*\geq 0 \qquad i=1,2,...,N.\end{aligned} \quad (17)$$

To solve the constrained optimization problem, the following primal Lagrangian form is obtained,

$$\begin{aligned}L\left(w,b,\xi,\xi^*,\alpha,\alpha^*,\beta_i,\beta_i^*\right)=\frac{1}{2}\|w\|^2+C\left(\sum_{i=1}^{N}\left(\xi_i+\xi_i^*\right)\right)-\sum_{i=1}^{N}\alpha_i\left(w_i\phi\left\{x_i\right\}+b-d_i+\varepsilon+\xi_i\right)\\ -\sum_{i=1}^{N}\alpha_i^*\left(d_i-w_i\phi\left\{x_i\right\}-b+\varepsilon+\xi_i^*\right)-\sum_{i=1}^{N}\left(\beta_i\xi_i+\beta_i^*\xi_i^*\right)\end{aligned} \quad (18)$$

The Lagrangian given in Equation (18) is minimized with respect to primal variables w, b, ξ, ξ^*; and, maximized with respect to nonnegative Lagrangian multipliers $\alpha,\alpha^*,\beta_i,\beta_i^*$. If Kuhn-Tucker conditions are used in the regression and in the regularized risk function given in Equation (14), we obtain the dual Lagrangian as,

$$J\left(\alpha,\alpha^*\right)=\sum_{j=1}^{N}d_i\left(\alpha_i-\alpha_i^*\right)-\varepsilon\sum_{j=1}^{N}\left(\alpha_i+\alpha_i^*\right)$$

$$-\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\left(\alpha_i-\alpha_i^*\right)\left(\alpha_j-\alpha_j^*\right)K\left(x,x_t\right) \tag{19}$$

which is subject to the constraint,

$$\sum_{j=1}^{N}d_i\left(\alpha_i-\alpha_i^*\right)=0, \qquad 0\le\alpha_i,\alpha_i^*\le C, \qquad i=1,2,...,N. \tag{20}$$

For the solution, first step is the calculation of α,α^* Lagrangian multipliers. In the second step, solving for w, the optimal weights of the regression is calculated:

$$w^*=\sum_{j=1}^{N}d_i\left(\alpha_i-\alpha_i^*\right)K\left(x,x_t\right). \tag{21}$$

Finally, the regression function is represented as,

$$f\left(x,\alpha_i,\alpha_i^*\right)=\sum_{j=1}^{N}d_i\left(\alpha_i-\alpha_i^*\right)K\left(x_i,x_j\right)+b^* \tag{22}$$

We have worked on SVM model with $K\left(x_i,x_j\right)$, whose value is the inner product of two x vectors in the ϕ feature space, $K\left(x_i,x_j\right)=\phi\left(x_i\right)\times\phi\left(x_j\right)$. Vapnik (1995) shows that any symmetric positive and semi-definite function that satisfies Mercer's conditions could be used as the Kernel functions. In SVM literature, common Kernel functions utilized are Gaussian, polynomial and linear kernel functions given below:

Gaussian: $$K\left(x_i,x_j\right)=\exp\left(\left(\frac{-\left\|x_i-x_j\right\|}{2\sigma^2}\right)^2\right) \tag{23}$$

Polynomial: $$K\left(x_i,x_j\right)=\left(x_i^T x_j+1\right)^d \tag{24}$$

Linear: $$K\left(x_i,x_j\right)=x_i^T x_j \tag{25}$$

In the study, we follow Ou and Wang (2010) and Perez-Cruz et. al. (2003) in the modeling process. We utilized Support Vector Regression models with Gaussian Kernel functions. For forecasting purposes, the researchers should follow several techniques such as risk minimization in the validation tests for selecting the kernel functions. The results are based on the data analyzed in the study. As a result, for different data sets, the researcher should consider applying different kernel functions to formulate the input space into the feature space to obtain well generalization performance. The selection of parameters in the kernel functions are discussed in many studies. Some methods include bootstapping and risk criteria minimization. Training of SVM regression is based on neural nets (Chang and Tsai, 2009), hybridization of SVM regression models with certain time series models (Chen et.al., 2008). Müller et.al. (1997) provide a comparison of models with different loss functions; namely, Huber's loss function and e-insensitive loss function. In the study, the SVR-GARCH model is restricted to be estimated with the Gaussian kernel function and loss function is defined as the e-insensitive loss function following Vapnik (1995). The hybrid modeling methodology followed in the study for the SVR-GARCH models is similar to Ou and Wang (2010) and Perez-Cruz et al. (2003) modeling procedure and follows the methodology derived for the NN-GARCH modeling (Donaldson and Kamstra, 1997).

Empirical results are given in Part 3. GARCH model estimated with ML will be compared with MLP-GARCH and SVR-GARCH models in terms of error criteria (MSE, RMSE) and Diebold Mariano (DM) predictive accuracy tests in forecasting.

At the first stage, we estimated SVR models for the conditional mean. At the second stage, the error terms are modeled with GARCH, FIGARCH, APGARCH and FIAPGARCH processes.

4 Data and Empirical Results

4.1 Data

Istanbul Stock Exchange ISE100 stock index data is gathered from the EVDS system of Central Bank of Turkey. The stock returns are calculated by using the daily closing prices of ISE100 covering the 11.02.1987-15.09.2011 period corresponding to 5957 observations. The stock returns data is calculated as first differenced logarithmic series, $y=ln(P_t/P_{t-1})$, where $ln(.)$ is the natural logarithms. In the process of SVR and MLP model estimation, the sample is divided into three sub-samples; namely, training, test and out-of-sample. Initial 80% of the observations are taken as the training sample. Following 20% of the observations corresponds to the test sample. The final 80 observations are left for out of sample forecasting. The daily returns are given in Figure 1. Areas in white correspond to the training sample; whereas, areas in grey and blue are the test and forecasting samples, respectively.

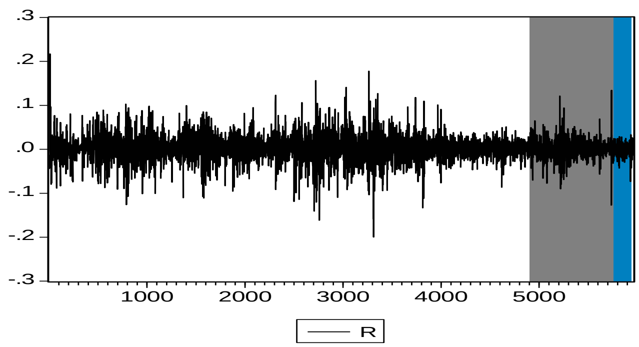

Figure 1. ISE100 Stock Index Daily Returns

Note. Areas in white: training sample (80% of obs.). Areas shaded in grey: test sample (20% of obs). Areas shaded in blue: out of sample set (last 80 obs.).

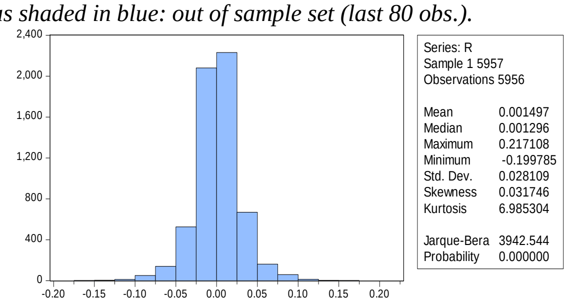

Figure 2. Histogram and Basic Statistics, ISE100 Daily Returns

Basic statistics are given in Figure 2. The mean of the series is 0.00149. JB normality test statistic is calculated as 3942.544 (*p*-value= 0.00000) and suggests that the null hypothesis of normality is strongly rejected. Skewness is calculated as 0.031746 and is close to zero, whereas the kurtosis is calculated as 6.985304 pointing to the leptokurtic structure, a phenomenon commonly observed for daily financial time series. ADF, PP and KPSS test statistics are calculated as -69.49 (0.0001), -69.77 (0.0001) and 0.31 showing that the daily return series is stationary[1]. The ARCH-LM(1-2) and ARCH-LM(1-5) test statistics are calculated as 434.85 (0.0000) and 138.69 (0.0000) and show that the data exhibits strong ARCH effects.

[1] ADF, PP unit root tests and KPSS stationarity test results are not given in detail to save space and could be obtained from the authors.

4.2 Econometric Results

At the first stage, among the GARCH family models, we selected basic GARCH model, and APGARCH models FIGARCH, taken as baseline models are estimated for evaluation purposes. Results are given in Table 1.

Table 1 Baseline Models

Coefficients:	***GARCH***	***APGARCH***	***FIGARCH***	***FIAPGARCH***
Cst(M)	0.0016** (5.36)	0.0014** (4.917)	0.0016** (5.488)	0.00145** (5.16)
Cst(V)	0.1998** (3.21)	0.7554 (1.20)	0.2842** (3.15)	0.15264 (1.08)
d-Figarch			0.4064** (8.21)	0.3797** (6.66)
ARCH	0.1572** (6.36)	0.1643** (7.07)	0.2317** (2.24)	0.2067* (1.76)
GARCH	0.8285** (29.99)	0.8326** (31.71)	0.4338** (3.73)	0.3793** (2.72)
APARCH (Gamma1)		0.0518* (1.64)		0.0592* (1.78)
APARCH (Delta)		1.6594** (8.18)		2.0994** (17.12)
LogL	13361.16	13366.35	13405.95	13410.05
AIC:	-4.5455	-4.5466	-4.5605	-4.5612
SIC:	-4.5410	-4.5398	-4.5548	-4.5532
Q(5)	14.1874 [0.00]	17.6420 [0.00]	5.8026 [0.12]	5.29105 [0.15]
Q(10)	26.9742 [0.00]	30.9775 [0.00]	14.085 [0.07]	12.6800 [0.12]
SB:	0.38924 [0.69]	0.4958 [0.61]	0.6236 [0.53]	0.8391 [0.40]
ARCH (1-2):	4.4142 [0.012]	6.1645 [0.00]	0.8560 [0.42]	0.55257 [0.57]
ARCH (1-5):	2.7684 [0.02]	3.4571 [0.00]	1.1540 [0.32]	1.0520 [0.38]

LogL: Loglikelihood, Q(p): pth order autocorrelation test, SB: Sign bias test, ARCH(p): pth order ARCH-LM test.

Included models have different characteristics to be evaluated; namely, fractional integration, asymmetric power and fractionally integrated asymmetric power models, namely, GARCH, APGARCH, FIGARCH and FIAPGARCH models. The ARCH and GARCH parameters are statistically significant at 5% significance level and the summation of alpha + beta is equal to 0.98574 and less than 1. The log-likelihood for the model is calculated as 13361.16. The AIC and SIC information criteria are calculated as -4.5455 and -4.5410 respectively. The data is daily and the highest autocorrelation is observed at the 5th and 10th days. Q(5) and Q(10) statistics show that autocorrelation of order 5 and 10 cannot be rejected at 5% significance level. ARCH-LM tests of order 1, 2 and 5 show that ARCH effect cannot be rejected. ARCH-LM test could also show possible nonlinearity in the residuals. Sign bias test statistic (SB) is calculated as 0.389 and sign bias is rejected by evaluating the news impact curve.

It is observed that, all volatility models perform better than the FIAPGARCH model in light of Log Likelihood criteria. If AIC and SIC criteria are evaluated, the lowest AIC (-4.5612) is calculated for the FIAPGARCH model; whereas, the lowest SIC is calculated as -4.5548 for the FIGARCH model. The sum of ARCH and GARCH parameters is calculated as 0.9857 for the GARCH model and similarly is less than 1 for the APGARCH, FIGARCH and FIAPGARCH model. For the fractionally integrated models, the differentiation parameters are estimated as 0.40 (FIGARCH) and 0.38 (FIAPGARCH).

Power terms obtained for returns calculated for stock indices in many developing economies are calculated comparatively higher than those obtained for the various indices in developed countries in various studies. The calculated power term is 1.65 in the APGARCH model and is estimated as 2.09 in the FIAPGARCH model showing high levels of asymmetry.

It is noteworthy to evaluate several studies. Power terms in single regime APGARCH models were calculated for daily returns as 1.57 in Nikkei 225 Index, as 1.81 in Hang Seng Index, as 1.69 in Kuala Lumpur Composite Index and as 2.41 in Singapore SES-ALL Index. The non-standard LR test is statistically significant and this suggests that linearity is strongly rejected. Further, STAR type nonlinearity tests are evaluated. Accordingly, the nonlinearity is accepted and linearity is rejected for the transition variable of one lagged daily returns.

It should be noted that, the GARCH model will be taken as the baseline model to be compared with the SVR-GARCH and MLP-GARCH models. Further, the study is restricted only GARCH type conditional volatility and aims to focus on the in sample and out of sample performance of the suggested hybrid versions, namely, SVR-

GARCH and MLP-GARCH architectures. MLP-GARCH model is estimated and results are given in Table 2.

Table 2. Architecture Selection and Training Results: MLP-GARCH Model

Model Type:*	**MLP-GARCH (4-6-1; 1,1)**
Training performance	0.103297
Test performance	0.110626
Overall performance	0.1069615
Training MSE	0.002
Test MSE	0.002246
Training algorithm	BP+CGD
Error function	SOS
Activation functions	Logistic
Output activation function	Identity

** Samples are divided into 80% training and 20 % test samples during optimization. Training/Test/Overall performance is shown with the Pearson's rho statistic. A MLP-GARCH (i,n,o; p,q) model is a model with i input variable, n neuron, o output variable with ARCH of order p and GARCH of order q. BP is back-propagation and CGD is the Conjugate Gradient Descent algorithm. SOS is the sum of squares.*

At the first step, lag length is selected with SIC information criteria as 4. Following the model selection procedure as discussed in the previous section, 100 models MLP-GARCH models are estimated with varying neurons in the hidden layer. The activation function is restricted to logistic activation function for its sigmoid shaped continuous function characteristics. The optimum MLP model is selected as a MLP-GARCH (4-6-1; 1,1) model with 4 input variable - 6 neurons in the hidden layer – 1 dependent variable in the output layer model with GARCH (1, 1) process. The output layer is restricted to be linear identity function which is a common practice in the NN literature. Training/Test/Overall performance is shown with the Pearson's rho statistic. A MLP-GARCH (*i,n,o; p,q*) model is a model with i input variable, n neuron, o dependent variable model with ARCH of order p and GARCH of order q. BP is back-propagation and CGD is the Conjugate Gradient Descent algorithms. SOS is sum of squares. Further, the estimated MLP-GARCH models are reported with their coefficients and diagnostics tests in Table 3. The ARCH and GARCH parameters are statistically significant at 5% significance level. The stationarity condition, alpha+beta<1 is achieved for all the MLP-GARCH models estimated. For a typical, alpha+beta is equal to 0.9850, hence, compared to the baseline GARCH model reported in Table 1 with alpha+beta = 0.9857, the results show improvement over the baseline GARCH, though the improvement is comparatively low.

Table 3 Multilayer Perceptron–GARCH Models, Parameter Estimates

	MLP-GARCH	***MLP-APGARCH***	***MLP – FIGARCH***	***MLP – FIAPGARCH***
Cst(M)	-0.0002 (-0.68)	-0.00001 (-0.06)	-0.0002 (-0.78)	-0.00003 (-0.12)
Cst(V)	0.2003** (3.25)	0.5808 (1.19)	0.0133 (1.29)	0.0153 (0.7)
d-Figarch			0.4332** (6.2)	0.4259** (7.45)
ARCH	0.1560** (6.37)	0.1634** (6.91)	0.1930** (1.99)	0.1848* (1.89)
GARCH	0.8290** (30.31)	0.8305** (31.08)	0.4290** (3.39)	0.4146** (3.26)
APARCH (Gamma1)		-0.060* (-1.82)		-0.0675** (-1.96)
APARCH (Delta)		1.7318** (8.52)		1.9684** (12.07)
LogL	13372.56	13377.5	13417.23	13421.37
AIC:	-4.5525	-4.5535	-4.5674	-4.5681
SIC:	-4.548	-4.5467	-4.5617	-4.5602
Q(5)	12.4986 [0.01]	14.3909 [0.00]	6.6036 [0.08]	6.2299 [0.10]
Q(10)	25.1365 [0.00]	26.9583 [0.00]	15.5264 [0.05]	14.8638 [0.06]
SB:	0.8317 [0.41]	0.8713 [0.3835]	1.0132 [0.31]	1.205 [0.22]
ARCH (1-2):	3.9319 [0.01]	4.8583 [0.01]	1.132 [0.32]	0.9353 [0.39]
ARCH (1-5):	2.4587 [0.03]	2.8385 [0.01]	1.3099 [0.26]	1.2395 [0.28]

*Note: p-values are given in brackets. t-statistics are given in parentheses. * (**) denotes %10 (%5) significance level.*

The ARCH and GARCH parameters are statistically significant at 5% significance level. The stationarity condition, alpha+beta<1 is achieved for all the MLP-GARCH models estimated. For a typical, alpha+beta is equal to 0.9850, hence, compared to the baseline GARCH model reported in Table 1 with alpha+beta = 0.9857, the results show improvement over the baseline GARCH, though the improvement is comparatively low. The AIC and SIC information criteria are calculated as -4.5525 and -4.548 which are both lower than the calculated AIC and SIC information criteria for the baseline GARCH model (AIC=-4.5455, -4.5410). Box-Pierce serial correlation statistic Q(5) is calculated as 12.4986 and serial correlation up to order 5 cannot be rejected. Q(10) statistic is calculated as 23.13 and show that serial correlation up to order 10 cannot be rejected. Further, ARCH-LM(1-2) and ARCH-LM(1-5) tests test remaining ARCH effects in the error terms up to order 2 and 5 and the results show that ARCH effect cannot be rejected for the MLP-GARCH and MLP-APGARCH models. It will be noted that, following the fractional integration specifications, both the MLP-FIGARCH and its asymmetric power version, the MLP-FIAPGARCH model show improvement in capturing both autocorrelation (based on Q tests) and volatility (based on ARCH-LM tests) compared to MLP-GARCH and MLP-APGARCH models. For a typical, the sign bias test statistic (SB) is calculated as 0.83 for the MLP-GARCH model and sign bias is rejected (p-value=0.41). Similarly, it could be concluded that, asymmetry is rejected for all of the MLP based models. As a result, the MLP-GARCH family captured asymmetry effectively for ISE100 daily returns series. The fractional differentiation parameters were estimated as 0.41 and 0.39 for the baseline models, FIGARCH and FIAPGARCH, respectively. The fractional differentiation parameter estimates showed an increase and were calculated as 0.43 for their MLP variants, the MLP-FIGARCH and MLP-FIAPGARCH models. As a result, the MLP augmented versions pointed at comparatively higher persistence than their baseline counterparts.

Following the methodology discussed above, the estimated SVR-GARCH models are reported in Table 4 and parameter estimates are given in Table 5. The Support Vector Regression variants of the MLP-GARCH models constitute four models, the SVR-GARCH, SVR-APGARCH, SVR-FIGARCH and SVR-FIAPGARCH; respectively. For the conditional mean process, the decision constant is calculated as -0.342241 the capacity coefficient is taken as 1 and the *nu* parameter is calculated as 0.4.

The gamma coefficient is taken as 0.25, the default value in the RBF kernel functions. MSE is calculated as 0.001. The SVR model is estimated with 1786 number of support vectors; whereas, the MSE and MAE for the process is calculated as 0.00082 and 0.0205 that insights good fit of the model to the ISE100 stock index daily returns data.

Table 4 SVR-GARCH Model Architecture

Support Vector Regression Architecture*	***Value:***
Decision Constant:	-0.342241
Capacity:	1
Nu:	0.4
Kernel type:	RBF (gamma=0.25)
No.of support vectors	1786 (1784 bounded)
MAE:	0.020560660
MSE:	0.000817462

** Samples are taken as: 80% training; 20 % test.*

The ARCH and GARCH parameters are statistically significant at 5% significance level. alpha+beta stationarity condition is achieved and is less than 1 for the four SVR based GARCH models. The AIC and SIC information criteria are calculated lower for the models analyzed. As a result, on overall conclusion could be made as the fact that, considering the baseline GARCH, MLP-GARCH and SVR-GARCH models, though the MLP-GARCH type models provided better goodness of fit compared to the baseline models, SVR-GARCH type models provided improvement in terms of fit. However, it should be noted that, the reported AIC and SIC type criteria constitutes to the in-sample forecasting capability; whereas for comparative purposes in forecasting, out-of-sample data forecasting results would provide better information. As a typical, for the variants of the FIAPGARCH models, SVR-FIAPGARCH provided the lowest (most negative) AIC and SIC statistics among all models (AIC=-4.5712, SIC=-4.5633). Box-Pierce serial correlation Q test statistics up to order 5 and 10 show that autocorrelation cannot be captured by the SVR-GARCH and SVR-APGARCH models, whereas, the Q(5) and Q(10) statistics showed that, SVR-FIGARCH and SVR-FIAPGARCH models, the fractionally integrated variants, show no autocorrelation in the error terms.

The result is similar to the findings discussed for he MLP-FIAPGARCH and MLP-FIGARCH models. ARCH-LM(1-2) and ARCH-LM(1-5) tests show that ARCH effect is strongly rejected for the SVR-FIGARCH and SVR-FIAPGARCH models. Further, similar to the MLP-GARCH models, sign bias tests show that sign bias is rejected for all SVR-GARCH family models and asymmetry is effectively captured.

Table 4 Support Vector Regression-GARCH Results

	SVR-GARCH	***SVR-APGARCH***	***SVR – FIGARCH***	***SVR – FIAPGARCH***
Cst(M)	-0.0007** (-2.43)	-0.0009** (-3.22)	-0.0006** (-2.48)	-0.0009** (-3.30)
Cst(V)	0.2029** (3.29)	0.5806 (1.186)	0.1441 (1.24)	0.1863 (0.74)
d-Figarch			0.4383** (5.88)	0.4246** (7.44)
ARCH	0.1562** (6.41)	0.1639** (6.89)	0.1804* (1.83)	0.1697* (1.67)
GARCH	0.8283** (30.33)	0.8282** (30.68)	0.4231** (3.15)	0.4009** (3.02)
APARCH (Gamma1)		0.0691* (1.94)		0.0793** (2.16)
APARCH (Delta)		1.7411** (8.58)		1.9373** (12.09)
LogL	13380.99	13386.35	13425.37	13430.57
AIC:	-4.5554	-4.5565	-4.5702	-4.5712
SIC:	-4.5509	-4.5497	-4.5645	-4.5633
Q(5)	7.5579 [0.18]	8.3849 [0.14]	5.5885 [0.34]	5.8589 [0.11]
Q(10)	23.5501 [0.01]	24.91 [0.005]	24.2028 [0.01]	14.183 [0.08]
SB:	0.3691 [0.71]	0.5119 [0.61]	0.5135 [0.60]	0.7163 [0.4738]
ARCH (1-2):	3.4654 [0.03]	4.1581 [0.02]	1.0521 [0.35]	0.9206 [0.39]
ARCH (1-5):	2.1908 [0.05]	2.4736 [0.03]	1.239 [0.29]	1.1677 [0.32]

*Note: P-values are given in brackets. t-statistics are given in parentheses. * (**) denotes %10 (%5) significance level.*

Since the fractionally integrated models showed significant improvement, the fractional differentiation parameters are evaluated with attention. It should be noted that, the *d* parameter estimates reported for the SVR-FIGARCH and SVR-FIAPGARCH models are 0.43 and 0.42 and show stationarity. Further, the parameter estimates are comparatively larger than the parameter estimates reported for the FIGARCH and FIAPGARCH baseline models. However, the *d* parameter estimates are lower than 0.5, hence, they point at no long memory in the processes.

In this section, the models will be evaluated for their generalization capabilities in the out-of-sample forecasting in terms of the RMSE error criteria. Results are given in Table 5 for 3 different model groups, the baseline, the MLP and their SVR variants.

Table 5 Out of Sample Forecast Statistics, 1-80 Working Days (4 Months) Ahead

	GARCH	**APGARCH**	**FIGARCH**	**FIAPGARCH**
RMSE:	0.000819000 [3rd]	0.00083000 [3rd]	0.00079600 [3rd]	0.00078900 [3rd]
	MLP-GARCH	**MLP-APGARCH**	**MLP-FIGARCH**	**MLP-FIAPGARCH**
RMSE:	0.000777800 [2nd]	0.000787800 [2nd]	0.000740200 [2nd]	0.000741200 [2nd]
	SVR-GARCH	**SVR-APGARCH**	**SVR-FIGARCH**	**SVR-FIAPGARCH**
RMSE:	0.000750000 [1st]	0.000761800 [1st]	0.000709900 [1st]	0.00071300 [1st]

**RMSE: Root Mean Squared Error, MAE: Mean Absolute Error*
**Models are ordered from the lowest error criteria (for both RMSE and MAE) to the highest. The rank of each model is given in [ ] brackets. Models are evaluated in terms of their capability in forecasting the conditional mean and variance separately.*

In Table 5, 12 different conditional volatility models are compared to investigate their forecast accuracy for 80 work days (4 month period) ahead. The root mean squared error criterion (RMSE) is reported at the first row and three models with the similar conditional variance processes is evaluated among its own group. The RMSE is calculated as 0.00082 for the baseline model; whereas the RMSE's of MLP-GARCH and SVR-GARCH models are calculated as 0.000778 and 0.00075, respectively. As a result, among the GARCH models, the baseline model has the 3rd rank, MLP-GARCH model shows improvement over the baseline GARCH and takes the 2nd place. Further, among the GARCH processes, the lowest RMSE is achieved for the SVR-GARCH model and the model achieves the 1st place among the models with GARCH type conditional volatility. Among the APGARCH type models, a similar result is obtained. The APGARCH model also showed the worst performance with RMSE=0.0083, however, its MLP and SVR based augmentations showed slight improvement with RMSE's equal to 0.00078 and 0.00076. The SVR-APGARCH model takes the 1st rank followed by the MLP-APGARCH model. For the FIGARCH and FIAPGARCH models, the in-sample results suggested that MLP and SVR augmentations showed significant improvement over other models estimated. The out-of sample analysis provided a similar conclusion for forecast accuracy. Though MLP-FIGARCH and MLP-FIAPGARCH models provided significant improvement over the baseline models, the SVR variants, SVR-FIGARCH (RMSE=0.0007099) and SVR-FIAPGARCH (RMSE=0.00071300) models showed the best performances among the fractional integration models and also compared to the MLP-GARCH models.

Lastly, the models are evaluated within the identical conditional mean processes. Among the baseline GARCH models the lowest RMSE (=0.000789) is achieved for the FIAPGARCH model which is followed by the FIGARCH model. Among the MLP-GARCH models, the lowest RMSE is achieved by the MLP-FIGARCH model (=0.000740) and is closely followed by the MLP-FIAPGARCH model (=0.000741). Among the SVR-GARCH models, the SVR-FIGARCH model provided the lowest RMSE (=0.000709) which suggests a 15.5% improvement over the simple GARCH and 4.5% improvement over the MLP-GARCH model. The SVR-FIGARCH model is closely followed by the SVR-FIAPGARCH model (RMSE=0.000713) which provided %14.8 improvement over the baseline GARCH model and 3.7% improvement over the MLP-FIGARCH model. The results suggest that, overall, SVR-GARCH and MLP-GARCH models benefit from their respective learning algorithms and provide better forecast accuracy compared to the simple GARCH model. On the other hand, the fractional integration and asymmetric power characteristics improve the forecast capacity even more. According to the results, SVR-GARCH and MLP-GARCH group models capture the volatility characteristics of Istanbul Stock Index daily returns, taken as an example of an emerging market time series with strong volatility characteristics.

5 Concluding Remarks

The study aimed to investigate linear GARCH, fractionally integrated FI-GARCH and Asymmetric Power APGARCH models and their nonlinear counterparts based on a family of Neural Networks and Support Vector Machhine models. In the study, nonlinear augmentations based on SVR and MLP models are taken as the basis. The SVR-GARCH models are proposed in spirit of NN-GARCH architectures. The models are tested for in-sample and out-of-sample forecasting the daily stock returns in ISE100 Istanbul stock index, an example of an emerging market with high volatility and leptokurtic distribution characteristics. The empirical results show that SVR-

GARCH models and NN-GARCH models provided significant gains compared to simple GARCH models and therefore the results suggested that the learning algorithms provided significant gains in forecasting.

The models are subject to modeling conditional mean processes with machine learning and neural network methodologies while the conditional variance processes are allowed to integrate fractional integration and asymmetric power characteristics. The future studies should focus on incorporating machine learning and neural network learning algorithms with GARCH models that also benefits from nonlinear econometric techniques such as regime switching and smooth transition regressions. The results are restricted to the data analyzed in the study similar to many empirical analyses. To obtain generalization, applications of the models to different financial markets are suggested for future research.

REFERENCES

[1] **Baillie, R.T., T. Bollerslev and H.O. Mikkelsen (1996),** ***Fractionally Integrated Generalized Autoregressive Conditional Heteroskedasticity*****;** *Journal of Econometrics* 74, 3-30;

[2] **Bildirici, M. and Ö. Ersin (2009),** ***Improving Forecasts of GARCH Family Models with the Artificial Neural Networks: An Application to the Daily Returns in Istanbul Stock Exchange***; *Expert Systems with Applications* 36, 7355-7362;

[3] **Bildirici, M. and Ö. Ersin (2013),** ***Forecasting Oil Prices: Smooth Transition and Neural Network Augmented GARCH Family Models*****;** *Journal of Petroleum Science and Engineering* 109, 230-240;

[4] **Bildirici, M. and Ö. Ersin (2014),** ***Modeling Markov Switching ARMA-GARCH Neural Networks Models and an Application to Forecasting Stock Returns*****;** *The Scientific World Journal* 2014, Article ID 497941, 1-21;

[5] **Bollerslev, T. (1986),** ***Generalized Autoregressive Conditional Heteroskedasticity*****;** *Journal of Econometrics*, 31(3), 307-327;

[6] **Cao, L. and Gu, Q. (2002),** ***Dynamic Support Vector Machines for Non-stationary Time Series Forecasting***; *Intelligent Data Analysis* 6, 67-83;

[7] **Chang, B., Tsai, H. (2009),** ***Training Support Vector Regression by Quantum-Neuron-Based Hopfield Neural Net with Nested Local Adiabatic Evolution;*** *International Journal of Innovative Computing, Information and Control* 5(4), 1013-1026;

[8] **Chen, S., Jeong, K., Haerdle, W. (2008),** ***Support Vector Regression Based GARCH Model with Application to Forecasting Volatility of Financial Returns*****;** *Economic Risk Discussion Papers*, SFB 649, 2008-14, 1-25;

[9] **Christianini,N. and Shaw-Taylor,J. (2000)**, ***An Introduction to Support Vector Machines and other Kernel-Based Learning Methods*,** 1st Edition; Cambridge: Cambridge University Press;

[10] **Cortes, C. and V. Vapnik (1995)**, ***Support-Vector Networks*;** *Machine Learning* 20, 273-297.

[11] **Cybenko, G. (1989),** ***Approximation by Superpositions of a Sigmoidal Function;*** *Math. Control Signals Systems* 2, 303-314;

[12] **Ding, Z., C. Granger, R. Engle (1993*), A Long Memory Property of Stock Market Returns and a New Model;*** *Journal of Empirical Finance* 1, 83-106;

[13] **Donaldson, R. and M. Kamstra (1996),** ***Forecast Combining with Neural Networks;*** *Journal of Forecasting* 15(1), 49-61;

[14] **Engle, R.F. (1982),** ***Autoregressive Conditional Heteroskedasticity with Estimates of the Variance of United Kingdom Inflation*;** *Econometrica* 50, 987-1007;

[15] **R.F. Engle (1990),** ***Discussion on Schwert*;** *The Review of Financial Studies* 3, 103–106;

[16] **Hagerud, G.E. (1997),** ***A Smooth Transition ARCH Model for Asset Returns*;** *SSE/EFI Working Paper Series in Economics and Finance* No.162;

[17] **Hamilton, J. D. and Susmel, R. (1994),** ***Autoregressive Conditional Heteroskedasticity and Changes in Regime*;** *Journal of Econometrics* 64, 307-333;

[18] **İnce, H., Trafalis, T. (2006),** ***A Hybrid Model for Exchange Rate Prediction,*** *Decision Support Systems* 42, 1054-1062;

[19] **Kara, Y., Boyacıoğlu, M., Baykan, Ö. (2011),** ***Predicting Direction of Stock Price Index Movement Using Artifical Neural Networks and Support tor Machines: The Sample of the Istanbul Stock Exchange*;** *Expert Systems with Applications* 38, 5311-5319;

[20] **Müller, K., Smola, A., Raetsch, G., Shoelkopf, B., Kohlmorgen, J., Vapnik, V. (1997),** ***Predicting Time Series with Support Vector Machines*;** *Proceedings of ICANN'97*, Springer, LNCS 1327, 999-1004;

[21] **Perez-Cruz, F., J. Afonso-Rodriguez and J. Giner (2003),** ***Estimating GARCH Models Using Support Vector Machines;*** *Quantitative Finance* 3, 1–10;

[22] **Rumelhart, D.E., Hinton, G.E., & Williams, R.J. (1986),** ***Learning Internal Representations by Backpropagating Errors****; Nature* 323, 533-536;

[23] **Ou P., H. Wang (2010),** ***Predicting GARCH, EGARCH and GJR Based Volatility by the Relevance Vector Machine: Evidence from the Hang Seng Index*;** *International Research Journal of Finance and Economics* 39, 46-63;

[24] **Özdemir, A., S. Tolun, E. Demirci (2011),** ***Endeks Getirisinin Yönünün İkili Sınıflandırma Yöntemiyle Tahmin Edilmesi: İMKB100 Endeksi Örneği*;** *Niğde Üniversitesi İİBF Dergisi* 4 (2), 45-59;

[25] **Vapnik, V. (1995),** ***The Nature of Statistical Learning Theory***, New York: Springer Verlag.