A RBF-EGARCH NEURAL NETWORK MODEL FOR TIME SERIES FORECASTING

ASRI BEKTIPRATIWI, MOHAMMAD ISA IRAWAN

Abstract. In this paper, we propose an alternative forecasting model based on the combinations between radial basis function model and exponential generalized autoregressive conditional model. We provide radial basis function to model the conditional mean and EGARCH to model the conditional volatility. We propose a regression approach to estimate the weights and the parameters of EGARCH using maximum likelihood estimator. The relevant explanatory variables are chosen based on its contribution of giving greater reduction in the in-sample forecast errors. We call this radial basis function model combined with EGARCH model as a RBF-EGARCH Neural Network Model. We use this model to provide forecasts of the stock returns.

Keywords and Phrases : Forecasting, Regression, RBF Neural Network, EGARCH .

1. INTRODUCTION

Time series models are well known for the future predictions. ARIMA model is one of the most widely used time series forecasting models in practice. In the conventional econometric models, the variance of the disturbance term is assumed to be constant. In such circumstance, the assumption of a constant variance (homoskedasticity) is inappropriate [4]. Engle [5] found evidence that for some kinds of macroeconomic data, the disturbance variances in time series models were less stable than usually assumed. He suggested the autoregressive conditionally heteroscedastic (ARCH) model as an alternative to the usual time series process to model the volatility of stock return. Later Bollerslev [2] generalized the model as GARCH to capture a higher order of ARCH. Nelson [8] proposed an extended version of such models: Exponential GARCH. In contrast to the conventional GARCH specification, which is requires non-negative coefficient, the EGARCH model does not impose non-negative constraints on the parameter space since it models the logarithm of the conditional variance. EGARCH has proven to be ideal to model the volatility of stock returns.

Artificial neural networks, on the other hand, have been successfully applied to signature recognition, classification analysis, pattern recognition problems and also time series forecasting. Radial basis function neural networks (RBF) have been extensively studied by researchers in time series forecasting area [9,10]. RBF is the major class of neural network model in which the activation function of a hidden unit is determined by the distance between the
input vector and a prototype vectors (usually called as basis function centers). We use clustering techniques, discussed in Subsection 1.2.2, to find a set of centers which more accurately reflects the distribution of the data points. Coelho and Santos [3] proved that the best performance in terms of point forecasts was achieved when the clustering algorithm k-means was employed.

In this study, we aim to extend the traditional RBF neural network model by combining with the EGARCH to model the variability of stock returns. We call the RBF neural network model combined with a EGARCH model for the residuals as RBF-EGARCH Neural Network model. We proposed a regression approach to estimate the weight and the parameters of EGARCH. The previous study by Coelho and Santos [3] proved that the combination of radial basis function neural network and GARCH (called as RBF-NN-GARCH) is able to provide accurate forecast. They used a joint estimation of RBF-NN and GARCH parameters via maximum likelihood along with the genetic algorithm to maximize the likelihood function.

This paper is organized as follows. We begin in Subsection 1 with description of the proposed model and the forecast methodology. Section 2 simulation result of its application to stock returns. Finally, Section 3 gives some concluding comments.

1.1 Proposed Model

Radial Basis Functions (RBF) networks are able to provide a local representation of a multi-dimensional space. Consider a mapping from a \( m \)-dimensional input space \( x \) to a \( n \)-dimensional target space \( y \). The data set consists of input vectors \( x \in \mathbb{R}^m \) together with corresponding target \( y \in \mathbb{R}^n \). The RBF generally uses a linear transformation for the output units and a nonlinear transformation function (basis function) as an activation function in the hidden layer. Fig. 1 shows a general structure of RBF neural networks. Let \( z_t \) be a vector which contains explanatory variables, the output of the mapping is then taken to be a linear combination of the basis functions

\[
y_t = \sum_{i=1}^{n} w_i \varphi(||z_t - \mu_i||)
\]

where \( w_i \) are the weights, \( \varphi(\cdot) \) is the basis function and \( \mu_i \) is the center of the \( i \)th basis function. We take a Gaussian radial basis function as the basis function where the width parameters \( \sigma_i \) control the amount of overlap of the radial basis function as well as the network generalization [6]. The Gaussian basis function is given by:

\[
\varphi_i(x) = \exp \left( -\frac{x^2}{2\sigma_i^2} \right)
\]

In the financial time series, it is often important to model a time-varying volatility. Volatility leverage effects, the observation that bad news has a larger impact on volatility than does good news, is well documented in the finance literature. The EGARCH of Nelson [8] has proven to be ideal for capturing the stylized facts that define stock return volatility, namely volatility clustering. In this sense, modeling time-varying volatility can improve the performance of our forecasting model. In this paper we consider the following extension of RBF neural network model (1). We combined the RBF neural network with a
EGARCH model for the residuals which is called as RBF-EGARCH Neural Network model. The RBF-EGARCH Neural Network model with \( n \) hidden units is specified by

\[
y_t = \sum_{i=1}^{n} w_i \varphi(||z_t - \mu||) + \varepsilon_t
\]

\[
\varepsilon_t = u_t \sqrt{h_t},
\]

\[
\ln h_t = \omega + \beta \ln h_{t-1} + \alpha \frac{|\varepsilon_{t-1}|}{\sqrt{h_{t-1}}} + \delta \frac{\varepsilon_{t-1}}{\sqrt{h_{t-1}}}
\]

where the function \( \varphi(\cdot) \) is defined as (2) and \( u_t \sim IID(0,1) \). \( z_t \) is a vector of lagged explanatory variables, i.e., \( z_t = (y_{t-1}, y_{t-2}, \ldots, y_{t-k}) \). \( h_t \) is the estimate of the conditional volatility at time \( t \) and \( \omega, \beta, \alpha \) and \( \delta \) are parameters to be estimated.

As it is clear from the model specification, we allow for EGARCH (1,1) process in the conditional variance. We restrict attention to the first-order case as it is the most frequently used in practical applications.

### 1.2 Forecasting Using RBF-EGARCH Neural Network

#### 1.2.1 Input Selections

In the initial stage of modeling, the selection of the model inputs (lagged variables) is based on the inspection of the sample autocorrelation and partial autocorrelation. In the second stage of modeling, an iterative procedure is applied to choose the relevant explanatory variables [3]. This iterative procedure works as follows. First, the model is estimated and evaluated using only the one lagged value \( y_{t-1} \). Second, an additional lagged \( y_{t-h} \) where \( h = 2 \) is included. Third, the model is re-estimated and the reduction in the in-sample forecast error in comparison to the model estimated in the first step is computed. Fourth, the second and the third steps repeated until the selected lag is reached. Finally, the selected lagged variables are those that contributed to greater reductions in the in-sample forecast error.

![General structure of the RBF Neural Network](image-url)
1.2.2 K-means clustering method

The clustering method used in this application is the classical k-means method. We use K-means clustering method to obtain the centers of each hidden unit. Given \( m \) data \( x_1, x_2, \ldots, x_m \), its implementation follows steps below:

Step 1: Select the number \( l < m \) of clusters
Step 2: Take the first \( l \) learning data \( x_1, x_2, \ldots, x_l \) as the center vectors:
\[
\mu_j = x_j, \quad j = 1, 2, \ldots, l
\]  
(4)
Step 3: Assign \( x_i \) (\( i = l + 1, l + 2, \ldots, m \)) to one of the clusters with the least distance criterion; that is, \( x_i \) belongs to the \( j \)th cluster if
\[
||x_i - \mu_j|| = \min_j ||x_i - \mu_j||, \quad 1 \leq j \leq l
\]  
(5)
Step 4: Recompute the center vectors using the new mean, that is
\[
\mu_j = \frac{1}{m_j} \sum_{i \in \mu_j} x_i, \quad 1 \leq j \leq l
\]  
(6)
where \( m_j \) is the number of the learning data belonging to the \( j \)th cluster \( \mu_j \).
Step 5: Assign \( x_i \) (\( i = 1, 2, \ldots, n \)) to one of the clusters with the nearest distance criterion.
Step 6: If at least one data point switch to another cluster, then recomputed the centers using the new mean and go to Step 5; otherwise, stop the procedure.

1.2.3 Model Estimation

The estimation of weight and EGARCH parameters is estimated via regression approach; see Medeiros et al. [7] for neural network regression with conditional volatility. As the true distribution of \( u_t \) is unknown, the parameters of model (3) are estimated by the maximum likelihood estimator. The MLE is usually referred to as a quasi maximum likelihood estimate (QMLE). Assume that the conditional distribution \( f(y_t | y_{t-1}, \ldots, y_1, \theta) \) is normal with mean \( \hat{y}_t \) and variance \( \hat{h}_t \). Define the true parameter vector of RBF-EGARCH neural network model with Gaussian function as \( \theta = (\theta_M, \theta_V)' \), where the conditional mean parameter vector is defined as \( \theta_M = (w_1, \ldots, w_n, \mu_1, \ldots, \mu_n, \sigma_1, \ldots, \sigma_n)' \) and \( \theta_V = (\omega, \beta, \alpha, \delta)' \) is the parameter vector for the conditional variance. Then, the likelihood function is
\[
f(y_1, \ldots, y_T; \theta) = \frac{1}{\sqrt{2\pi}} f(r_1; \theta) \prod_{t=2}^{T} \frac{1}{\sqrt{h_t}} \exp \left[ -\frac{1}{2} \frac{(r_t - \hat{y}_t)^2}{h_t} \right]
\]  
(7)
The MLE is obtained by maximizing the log likelihood function,
\[
\ln f(y_1, \ldots, y_T; \theta) = -\frac{T}{2} \ln (2\pi) + \ln f(r_1; \theta) - \frac{1}{2} \sum_{t=2}^{T} \ln (h_t) - \frac{1}{2} \sum_{t=2}^{T} \frac{(r_t - \hat{y}_t)^2}{h_t}
\]  
(8)
We can rewrite the first equation of model (3) as regression model,
\[
y = Z(\varnothing)w + \varepsilon
\]  
(9)
where \( y = (y_1', y_2', \ldots, y_T')' \), \( \varepsilon = (\varepsilon_1, \varepsilon_2, \ldots, \varepsilon_T)' \), \( w = (w_1, \ldots, w_n)' \) and
\[
Z(\varnothing) = \begin{pmatrix} 
\varphi(\|z_1 - \mu_1\|) & \ldots & \varphi(\|z_1 - \mu_n\|) \\
\vdots & \ddots & \vdots \\
\varphi(\|z_T - \mu_1\|) & \ldots & \varphi(\|z_T - \mu_n\|)
\end{pmatrix}
\]
with \( \phi = (\mu_1, \ldots, \mu_m, \sigma_1, \ldots, \sigma_n) \); \( \sigma_t \) is the width of basis function. Assuming \( \phi \) fixed, the initial weight parameter vector of the model can be estimated analytically by

\[
\hat{w}_0 = (Z(\phi)'Z(\phi))^{-1}Z(\phi)'y
\]  

(10)

1.2.4 Forecast Evaluation

Besides using the usual Root Mean Square Error (RMSE), to evaluate whether the result of the model can be used as trading strategy, the estimated (forecasted) value with same sign as the observed (true) value is calculated by metric SIGN.

Function (SIGN):

\[
SIGN = \frac{1}{N} \sum_{i=1}^{N} \delta_t
\]  

(11)

where \( \delta_t = \begin{cases} 
1, & \text{if } y_t \hat{y}_t \geq 0 \\
0, & \text{if otherwise}
\end{cases} \)

2. SIMULATION RESULTS

Our sample is the stock returns of Bank Rakyat Indonesia Tbk. It is daily, for the period 11 November 2003 – 11 March 2011, giving a total 1911 observations. The first 80% of the data was used to estimate the parameters of the model and the 20% of the data was used to provide out-of-sample forecasts. We evaluated the forecasting performance of the model based on the out-sample forecast. Fig. 2 shows the daily stock return and Table 1 presents some descriptive statistics.

Fig 2. Daily Return of Bank Rakyat Indonesia Tbk
The relevant explanatory variables are chosen based on the iterative procedure presented in Subsection 1.2.1. All those steps are repeated until lag $y_{200}$ are reached [3]. The following inputs that contributed to greater reduction are $z_t = y_{t-1}, y_{t-2}, y_{t-3}, y_{t-4}, y_{t-5}, y_{t-6}, y_{t-14}, y_{t-15}, y_{t-16}, y_{t-23}, y_{t-24}, y_{t-30}$. Fig 3 explains the sample ACF and PACF until lags 12. We find that the estimated optimal numbers of hidden units based on trial and error are 5 hidden units. The estimated width is 0.0221. The estimated optimal parameters of RBF-EGARCH Neural Networks model with 5 hidden units are shown in Table 2 and Table 3, respectively.

We compute daily point forecast from one- to 100-steps-ahead. Table 4 reports the forecasting result for RBF-EGARCH Neural Network model, ARMA(1,1)-EGARCH(1,1) and RBF-GARCH Neural Network model. The best performance based on RMSE is achieved by the ARMA(1,1)-EGARCH model. On the other hand, based on SIGN the best performance is achieved by the RBF-EGARCH Neural Network. It gives accuracy 63%. Finally, Fig 4 shows the forecasted conditional volatilities obtained with the RBF-EGARCH Neural Networks model for 300-step-ahead.
Table 4. Forecast Evaluated

<table>
<thead>
<tr>
<th></th>
<th>ARMA-EGARCH</th>
<th>RBF-GARCH Neural Network</th>
<th>RBF-EGARCH Neural Network</th>
</tr>
</thead>
<tbody>
<tr>
<td>RMSE</td>
<td>0.0081</td>
<td>0.0155</td>
<td>0.014</td>
</tr>
<tr>
<td>SIGN</td>
<td>0.59</td>
<td>0.60</td>
<td>0.63</td>
</tr>
</tbody>
</table>

3. CONCLUSION

In this paper we propose an alternative model for modeling the volatility of the conditional variances: A RBF-EGARCH Neural Networks Model. Our proposed forecasting model combines a RBF neural network for the conditional mean and a parametric EGARCH model for the conditional volatility. The regression approach is used to estimate the weight and the parameters of EGARCH model. Our simulation result based on sample of Bank Rakyat Indonesia Tbk stock returns indicate that our propose model is able to accurately predict 63% upward and downward movements of future predictions. The simulation results obtained in the forecasting performances motivate further work, which will involve comparing different method of parameters model estimation.
References


Asri Bektipratwi: Department of Mathematics, Institut Teknologi Sepuluh Nopember of Surabaya, Indonesia
E-mails: asri.bekti@gmail.com

Mohammad Isa Irawan: Department of Mathematics, Institut Teknologi Sepuluh Nopember of Surabaya, Indonesia
E-mails: m.isa.irawan@gmail.com