Neurocomputing 182 (2016) 197–209

Contents lists available at ScienceDirect

Neurocomputing

journal homepage: www.elsevier.com/locate/neucom

A new hybrid quadratic regression and cascade forward backpropagation neural network

Augustine Pwasong *, Saratha Sathasivam

School of Mathematical Sciences, Universiti Sains Malaysia, 11800 Pulau Pinang, Malaysia

ARTICLE INFO

Article history:
Received 7 August 2015
Received in revised form
3 December 2015
Accepted 12 December 2015
Communicated by G. Thimm
Available online 31 December 2015

Keywords:
Bayesian model averaging
Regression
Neural network
Mean absolute error
Root mean square error
Forecasting

ABSTRACT

In this study, a quadratic regression model (QRM) and a cascade forward backpropagation neural network (CFBN) are jointly integrated together to form a hybrid model called the new hybrid quadratic regression method and cascade forward backpropagation neural (QRM-CFBN) network method. The new hybrid method was tested on a daily time series data obtained from the UCI repository data link and the data set was collected from a combined cycle power plant. The joint integration was made possible by the Bayesian model averaging technique, which was used to obtain a combined forecast from the two separate methods. The model resulting from the joint integration was applied on the log difference series of the original time series data. The results obtained from the new hybrid QRM-CFBN were compared with the results obtained from the hybrid ARIMA-RNN, standalone cascade forward backpropagation neural (CFBN) network and layered recurrent neural network (LRNN) after being tested on the same sample time series data respectively. The comparison indicates that the results emerging from the new hybrid QRM-CFBN method on the average, generally results in better performance when compared with the hybrid ARIMA-RNN, the standalone CFBN and the standalone LRNN for 1 day, 3 days as well as 5 days prediction mean absolute error (MAE) and root mean square error (RMSE) for varying data samples of 50, 100, 200, 400 and 800 days respectively. The RMSEs and the MAEs were applied to ascertain the assertion that the new jointly integrated forecast has better forecasting performance greater than the standalone CFBN and LRNN forecasts as well as ARIMA-RNN forecast. The analysis for this study was simulated using MATLAB software, version 8.03.

1. Introduction

In this study, we propose an integrated approach to time series data forecasting, where we integrated a nonlinear cascade forward backpropagation neural network method with a linear quadratic regression method jointly. The intention for the joint integration between the quadratic regression method and the cascade forward backpropagation neural network method is to form a hybrid model that could capture diverse forms of relationship in a time series data. The new hybrid quadratic regression method and the cascade forward backpropagation neural network model proposed here takes advantage of the distinctive strength of regression models in time series forecasting and artificial neural networks in linear and nonlinear modeling. In modeling complex systems that possessed linear and nonlinear correlation structures, the joint integrated method that gives birth to a new hybrid model constitutes an effective fashion to improve forecasting performance in the mean square error sense.

Bates and Granger [1] first presented the hybrid forecasts by combining separate sets of forecasts of airline passenger data to form a hybrid set of forecast. The hybrid set of forecasts yields lower mean square error than either of the original forecast. Their combined forecasts were measured as a fruitful substitute to individual forecasting methods. Hybrid methods by combining regression methods were first carried out by Stone [2], followed by LeBlanc and Tibshirani [3], as well as Mojirsheibani [4]. Hybrid methods by combining time series methods were earlier carried out by Clemen [5], Weigend et al. [6], as well as Hansen and Nelson [7]. The entirety of these investigations advocated that hybrid methods had improved forecasting performance equated to individual forecasting methods. Wolport [8], Smith et al., [9] as well as Zhang and Beradi [10] were some of the earliest authors to propose hybrid methods by merging different neural network architectures and showed that hybrid methods involving different neural network architecture produced better forecasting performance compared to individual neural network methods. In our work, new trends in hybrid methods involving the combination of

* Corresponding author.
E-mail address: davougus@gmail.com (A. Pwasong).

http://dx.doi.org/10.1016/j.neucom.2015.12.034

classical statistical methods such as quadratic regression methods with different neural network architectures such as the cascade forward back propagation neural network are proposed.

Wedding et al. [11] describe a method of merging the forecasts from the Univariant Box–Jenkins (UBJ) method and the Radial Basis Function (RBF) networks, by means of certainty factor values resulting from the output of the RBF's basis functions to form a hybrid forecasting model, such that the resulting hybrid forecasting model produces more accurate outcomes than using RBF or UBJ forecasting alone. The drawback in this hybrid forecasting model is that the author employed a winner take all approach to be able to merge forecasts obtained from RBF network and UBJ method. There is no clear dichotomy combination of the two forecasts in this method, rather the user specifies a certainty cutoff point such that if the RBF's output certainty is greater than or equal to the cutoff value, then the RBF's forecast is used and if the RBF's certainty is less than the cutoff value, then the forecast from the UBJ model is used. This situation disregards the forecast of other models that may produce a good hybrid forecasts value. In our study of the new hybrid QRM-CFBN, the Bayesian modeling technique (BMA) is used to address the drawback of the UBJ–RBF model because it is a multi-model combination technique where it produces a combined forecast from two separate forecasts. in this study our two separate forecasts are forecast by the quadratic regression model (QRM) and forecast by the cascade forward backpropagation neural (CFBN) network.

Bianco et al. [12] developed a system of different regression models to estimate GDP, price and GDP per capita elasticity of domestic and non-domestic electricity consumption in Italy which they used to forecast the future growth of these consumptions. They used the ordinary least squares method to estimate the parameters of the different regression models, but it was not possible to conduct conventional inference, such as the *t*-test, hence, this trend created a major disadvantage in their model, due to the impossibility to confirm if the estimated coefficients are significantly different from zero. However, on comparing their model with other models for forecasting electricity in Italy such as the results presented by Gori and Takanen [13] and two national forecasts published by Terna [14] and CESI [15] it shows that there is a substantial agreement between the available national forecasts and the regression equations proposed by Bianco et al. much better than the other models. Another important feature of the proposed regression equations of Bianco et al. is that they are based on simple models which required only fundamental data as input, allowing cutting the cost linked to data mining which is one of the fundamental requirements for an econometric model. Our hybrid method intents to address the drawbacks of regression equations by Bianco et al. in order to avoid the inconvenience about the impossibility to conduct conventional inference on the coefficients by transforming our data using logarithmic transformation in order to make them stationary, thereby obtaining consistent forecasting outcomes.

Jie Wang and Jun Wang [16], in their study, applied multilayer perceptron (MLP) with a backpropagation algorithm and a stochastic time strength function to develop a stock price volatility forecasting model. They also presented a better technique which integrates the principal component analysis (PCA) into a stochastic time strength neural network (STNN) for predicting financial time series, called the PCA-STNN model. The PCA-STNN approach removes the principal components from the input data by the PCA technique, and used them as the input of the STNN method. This process can eradicate redundancies of the original data points and eliminate the correlation amid the inputs. Their study requires that the STNN input variables should have poor correlation, since the stout correlation amid input variables infers that they convey additional repetitive data points that may intensify the computational complexity and decrease the forecast precision of the model. The PCA-STNN method is not very effective when the input variables have a strong correlation, that is, when the input variables convey more recurrent information. Hence, our new integrated hybrid QRM-CFBN technique models the input variables by the QRM regression method, such that the residuals of the regression are then used as the target variable in the followed cascade forward backpropagation neural (CFBN) network. The QRM-CFBN method will further reduce the strong correlation between input variables greater than the PCA-STNN; thereby decreasing the computational complexity and increasing the prediction accuracy of the new QRM-CFBN hybrid model.

The remaining part of this study is organized as follows. Section 2 summarizes regression models and neural network models in forecasting. Section 3 describes the new hybrid methodology which is an integration of a quadratic regression technique and a cascade forward back propagation network method. Section 4 presents simulation results and discussions on the outcome of the results that emerges from the empirical investigation. Section 5 reports concluding remarks and future work.

2. Regression and neural network methods in forecasting

In their study, Adhikari and Agrawal [17] propose a method for linearly merging time series forecasts that endeavor to determine the linking weights after examining their patterns in successive in-sample predicting trials. The performance of their proposed forecast integration algorithm principally hinge on the values of the parameters M, N, and the architecture of the proposed ANN model where M is the number of forecasting trials and N is the appropriate size of the validation set. A very big value of M increases the magnitude of the in-sample weight set and the time complexity of the algorithm. Accordingly, the weights from many training-validation pairs are involved which are far away from the real testing data set. They found that increasing the number of forecasting trials after a certain level practically lowers the precision, instead of improving it. In another development, if M is excessively small, then the integration becomes bias to the latest in-sample observations and may be insufficient. Nevertheless, there is no theoretical standard to find the precise number of predicting trials. From their experiments with different values of M, ranging from 25 to 100, they found that the appropriate number of the forecasting trials lies between 40 and 50 for all-time series and as such, they make M to be 50 so that the size of the in-sample weight data set w is 200 for all-time series.

Adhikari and Agrawal [17] further assert that the method of using only a single model has a quantity of unavoidable limitations, which stimulated conjoining forecasts as a profitable substitute. An amalgamation of forecasts is hinged on the ultimate justification that any standalone model is predisposed to vague specification and insufficiency, while some non-optimal models collectively can very meticulously estimate the authentic data generation process. However, a forecast amalgamation can be operative only when there is sizeable magnitude of multiplicities amid the standalone models. The foregoing fact stimulates the integration of the quadratic regression method and the cascade forward backpropagation method, which constitute the cardinal subject of this study. Nikolopoulos et al. [18] compared precision of multiple linear regressions in forecasting the effect of different TV programs on Greek TV viewers by a simple bivariate regression model, three diverse types of artificial neural network and three methods of nearest neighbor analysis. They established that multiple linear regressions perform reasonably unwell. They also instituted that a greater accuracy was obtained from predictions hinged on a simple bivariate regression model, a simple nearest neighbor method and from two forms of artificial neural networks applied in their investigation.

Wang and Wang [16] proposed a neural network model called a stochastic time effective function neural network (STNN) with