Optimized Artificial Neural network models to time series

Marwan Abdul Hameed Ashour

Department of Statistics, College of Administration & Economics, University of Baghdad, Baghdad, Iraq.
E-mail address: dr_marwan@uobaghdad.edu.iq

Abstract:

Artificial Neural networks (ANN) are powerful and effective tools in time-series applications. The first aim of this paper is to diagnose better and more efficient ANN models (Back Propagation, Radial Basis Function Neural networks (RBF), and Recurrent neural networks) in solving the linear and nonlinear time-series behavior. The second aim is dealing with finding accurate estimators as the convergence sometimes is stack in the local minima. It is one of the problems that can bias the test of the robustness of the ANN in time series forecasting. To determine the best or the optimal ANN models, forecast Skill (SS) employed to measure the efficiency of the performance of ANN models. The mean square error and the absolute mean square error were also used to measure the accuracy of the estimation for methods used. The important result obtained in this paper is that the optimal neural network was the Backpropagation (BP) and Recurrent neural networks (RNN) to solve time series, whether linear, semilinear, or non-linear. Besides, the result proved that the inefficiency and inaccuracy (failure) of RBF in solving nonlinear time series. However, RBF shows good efficiency in the case of linear or semi-linear time series only. It overcomes the problem of local minimum. The results showed improvements in the modern methods for time series forecasting.

Keywords: Efficiency Neural Networks, Forecast Skill, Local Optimum, Optimization, Time Series.

Introduction:

Time series forecasting is a very powerful computational method that allows predicting, future outcomes of a system based on how the system previously acted, it has a wide range of applications in many scientific fields. Using neural networks and machine learning approaches can predict a wide range of different events, both of natural and human origin produced. Artificial neural networks (ANN) have received much attention in recent years, and the question of which type of neural networks are better in prediction has yet to be resolved. The ANN contains three-layer: input and output layer, hidden layer between input and output layers, and layers relate through the neurons (weights).

In statistics, forecasting the inputs for neural networks are typically the past observations of the info series and therefore the output is that the future value. Most lag structures are to manage the number of input units of the corresponding neural networks. The artificial neural cell consists of 4 parts are synapse, summation function, activation function, Axon paths. When estimating neural network models, it is possible to end up at a local optimum or not to converge to an optimum at all. We must use techniques to avoid these problems[1].

Artificial neural networks (ANN) are now widely accepted as a technology offering an alternative way to solve complex and ill-defined problems due to their strong non-linear mapping ability. Backpropagation neural network, radial basis function neural network (RBFNN), general regression neural network (GRNN), and recurrent neural network are some of the most common artificial neural network techniques (RNN).

The accuracy of three different approaches to time series prediction is compared in this study: recurrent neural networks, backpropagation neural networks, and radial basis function neural networks. The vital sorts of ANN are Backpropagation Network (BP), Radial Basis Function Neural networks (RBF), and Recurrent neural networks (RNN). The methods utilized in this paper's methodology are BP, RBF, and RNN. The efficiency of these networks in a linear and nonlinear model is also considered.

Previous studies have primarily focused on studying the difference between artificial neural networks, found that general regression neural network is better than backpropagation and radial basis function neural networks[2-4]. Other authors found that no significant difference between BP, RBF, and Elman[2,5,6]. In recent years, several studies have focused on using the hybrid models of artificial neural networks There were no significant differences between these models[7-11].

All previous studies concluded that all artificial neural network models are good and competitive, but they do not focus and study the ability of these models to avoid optimal local problems, which is one of the most important nonlinear estimation problems. The difference between the present work and previous work is to diagnose successful and efficient artificial neural network models to avoid optimal local problems and to identify models capable of solving both linear and nonlinear time series.

The objective of this paper is to determine whether models of artificial neural networks can avoid local optimum problems in time series forecasting. Also, identified which artificial neural networks are optimal for solving both linear and non-linear time series. on the practical part, it was used two sets of data; the first well-known real data set is that the energy data for Africa petroleum production from 1980 to 2012, represent a linear model and therefore, the second well-known real data set is for the Tunisian Dinar/Euro rate of exchange from January 2009 to February 2014, represent a non-linear model.

Materials and Methods:
Back Propagation (BP)

This network includes three-stage (Forward, Back Pass, Update Weights). During the forward propagation phase, the input signal propagates to each hidden level node, and then the activation value of each hidden layer node is determined for that signal. The signal of each of these nodes is then sent to each output layer once and then calculated to make the network response to the sample[8-10]. The activation value of each output layer node is calculated for that signal per node in the output layer. Activations are compared with the actual resulting value during the training process to determine the error value of that node The weights are updated based on the error derivative. Fig.1 illustrates the forward propagation of node activations from input nodes to output nodes and the backpropagation of in influence errors from output nodes to input nodes[11,12]
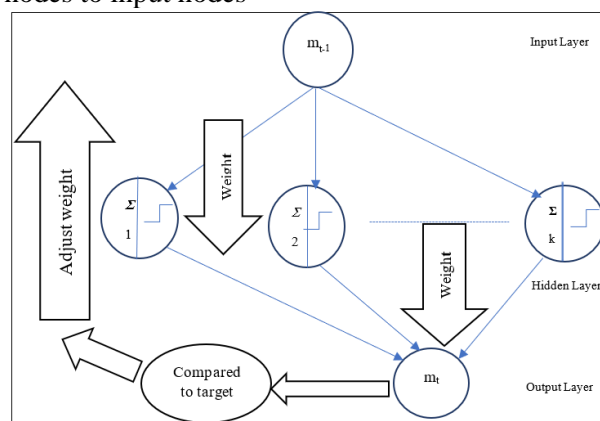

Figure1. BP methodology

General Regression Neural network

A general regression neural network is a type of RBF network that focuses on learning in a single pass. A GRNN is made up of RBF neurons in a hidden layer. Normally, the hidden layer contains a large number of neurons that serve as a training example. As the neuron's center and training example are associated, its output indicates how close the input vector is to the training example[13-15]. The GRNN architecture is shown in Fig. 2.[5,9].
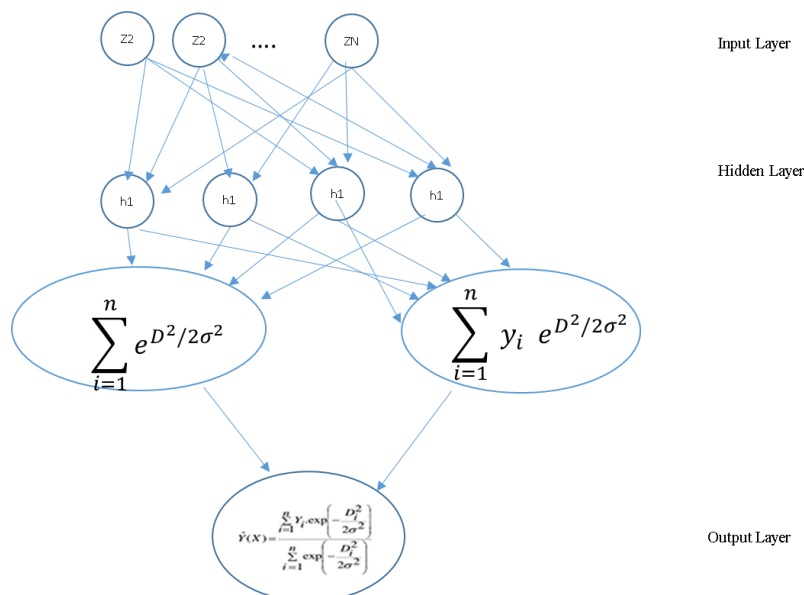
Figure 2. GRNN methodology.

Recurrent neural networks

RNN is a type of neural network that uses hidden units to analyze data streams. The output of some applications, such as text processing, speech recognition, and DNA sequences, is depending on previous calculations. RNNs are well suited for the health informatics domain, where enormous amounts of sequential data are available to process as they deal with sequential data[16-18]. Fig.3 depicts an RNN model [11,15].
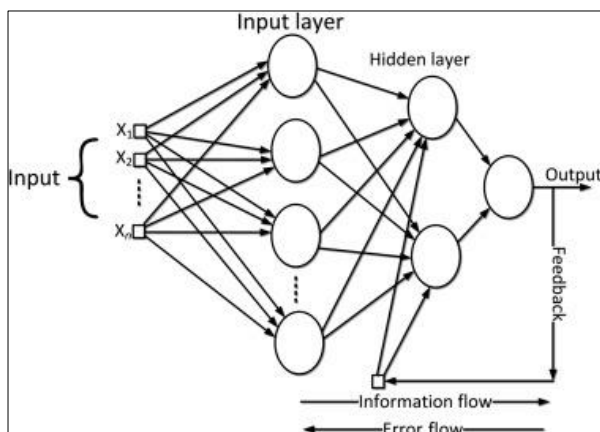

Figure 3. RNN structure.

Time series and Artificial Neural networks

ANN becomes more important in statistical processing and analyses. It's self-learning and training, due to its capacity. The major section of researchers uses shifting theorems supported or lagging statistics by one or more degrees to select the artificial neural input of a statistical network. The following formula chooses the neural network inputs[4,11].

$$z_t = w_0 + \sum_{j=1}^{q} w_j \cdot g\left(w_{0j} + \sum_{i=1}^{p} w_{ij}.z_{t-1}\right) + \varepsilon_t \qquad 1$$

Forecasting accuracy methods

To measure the accuracy and efficiency of the results of the methods used in this paper. The criteria were adopted (root square error (RMSE), mean absolute percentage (MAPE), and prediction skills (SS)[8].

$$RMSE = \sqrt[2]{\frac{\sum et}{n}} \qquad 2$$

$$MAPE = \frac{\sum\left|\frac{et}{Zt}\right|}{n} * 100 \qquad 3$$

Where z_t is actual value, z_t is forecast value, e_t is an error, $e_t = z_t - \hat{z}_t$, and n is a sample size.

Forecast skills are an extended representation of prediction error that compares the prediction accuracy of a given prediction model to another reference model, and are widely used in mathematical prediction as metrics for evaluating neural network performance[15]:

$$SS = 1 - \frac{MSE_f}{MSE_{Ref}} \qquad 4$$

$$MSE_f = \frac{1}{n}\sum_{i=1}^{n}(f_i - O_i)^2 \qquad 5$$

$$MSE_{Ref} = \frac{1}{n}\sum_{i=1}^{n}(\bar{O} - O_i)^2 \qquad 6$$

Where: O_i is observation values and f_i is forecast values.

If the SS value is 1, the prediction is perfect, and If SS is 0, negative skill[8,17].

Results and Discussion:

To diagnose the best types of neural networks for solving any time series, two well-known real data sets for two time series representing the different cases of the time series (nonlinear, linear,

semilinear) are used. Fig.4 depicts the first well-known real data set: energy data for Africa crude oil production from 1980 to 2012. The Tunisian Dinar/Euro exchange rate from January 2009 to February 2014 is the second well-known real data set, as shown in Fig.5.
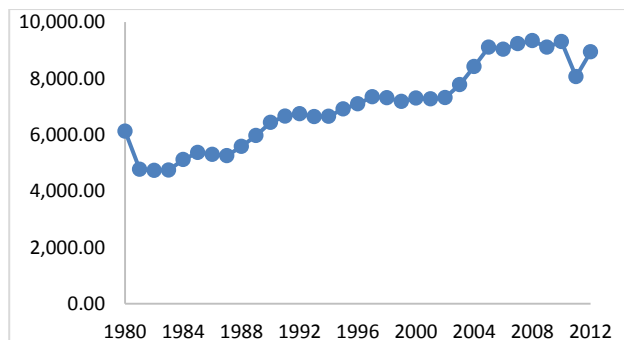

Figure 4. time-series of Africa oil production (Source: The United States Energy Information Administration)
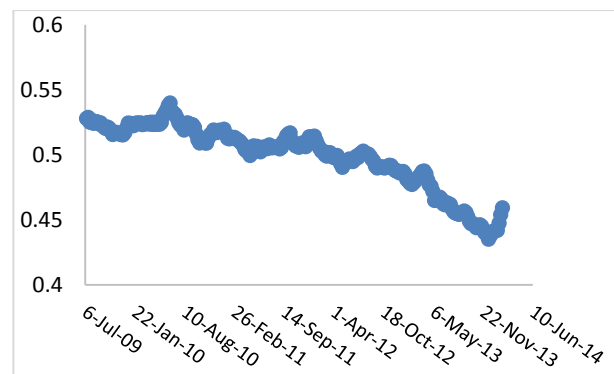

Figure 5. Time series of Exchange rate for Tunisian dinar/Euro (Source: Wall Street Journal).

The structure of ANN for two-time series as the following:

Input layer: the number of nodes in this layer is one node; represented by the variable (Z_{t-1}) based formula 1.

Hidden layer: the best number of nodes in this layer is 10 nodes for the first time series as the observing number of time series is small. There would be 25 nodes for the second time series as the observing number of time series is appropriate.

Output layer: This layer has one node, which is represented by the vector (Z_t). The MATLAB high-level programming language was used for the implementation. The quality performers of BP, ENN, and GRNN for the two-time series under study have been shown in Tab.1

It has shown from the evaluation of the quality of artificial neural network in Fig.6 the following:

In the first time series (linear series), a best linear fit that matches the target for all types of artificial neural networks (BP, ENN, and GRNN) Indicating the quality and efficiency of those networks.

In the second time series (nonlinear series), a best linear fit that matches the target for the BP and ENN indicating the quality and efficiency of those networks. Whereas the best linear fit of GRNN is very far from the target that indicates the inefficiency of that artificial neural network (GRNN).

The Result efficiency and accuracy error of three types ANNs for two time series-based formulas 2,5 and 6 had shown in Tab.1

Table1. Result efficiency& accuracy error of three types ANNs

Time series	Technique	Efficiency & Accuracy		
		RMSE	MAPE	SS
Linear series / Time series of Africa oil production	BP-1	3.97	0.015	1
	ENN-2	3.98	0.015	1
	GRNN-3	3.99	0.015	1
Nonlinear series (financial series)/ Time series of Exchange rate for Tunisian dinar/Euro	BP-4	0.0029	0.3	1
	ENN-5	0.0041	0.5	1
	GRNN-6	0.0265	4.21	0.0019

In the first time series (linear series), the error results in Table1 are shown the RMSE & MAPE for the three networks are equal approximated. In the second time series (nonlinear series), the error results in Tab.1 show that the minimized RMSE & MAPE are BP then ENN; the error results for all methods are relative to zero. They evaluate the efficiency of forecasting time series methods (ANNs types) as shown in Fig.6. The result of the efficiency of forecasting time series methods or the SS is indicated as follows:

In the first time series (linear series that all types of ANN are efficient. In a second time series (nonlinear series), that the BP and ENN are efficient only. While GRNN results are near zero, this indicates the inefficiency (negative skill) of that network.

In general, the results of the efficiency of ANNs types indicated that perfect forecasts for BP and ENN in linear, semilinear, nonlinear time series. Nonetheless, GRNN showed perfect forecast in linear time series only while the negative skill

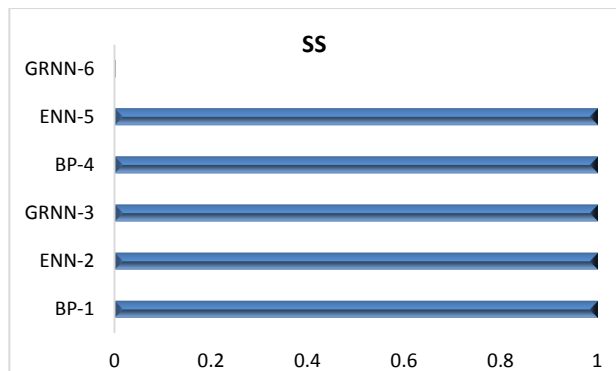forecast has shown for GRNN in nonlinear time series.

Figure 6. Histogram of SS

Thus, the worst method in its performance according to the efficiency and the accuracy was the GRNN-6.

The comparison results for the two-time series in Tab.1 show that the highest value of the MAPE resulted from the GRNN-6 method, which is congruous with the network assessment results in Tab.1 as well as the SS value. Moreover, the GRNN-3, BP-1, and ENN-2 give the highest value of the RMSE, which indicate that it is poor results according to this case. It does not match the network evaluation results in Table 1 and the SS value. Because the RMSE is affected by the second time-series observations, which are limited to 0.4 and 0.6, the error value is small and near to zero, whereas the MAPE value is affected.

Conclusions:

The results reveal that BP and ENN are better than GRNN. The results also show that inefficiency or inability (failure) of GRNN spatially, RBF generally to solve nonlinear time series and their efficiency to solve linear time series because it did not include (was without) feedback.
It was found that the efficiency of artificial neural network models to solve linear time series. The calculation in this work suggests that the optimal types of neural networks for solving linear, semi-linear, and nonlinear time series are BP and ENN. Its significance in terms of efficiency and accuracy. The result has proven the superiority of the BP and RNN in general, ENN spatially for avoiding local optimum problems, which is one of the very important problems of non-linear estimation. The results presented here may facilitate improvements in the time series forecasting modern method.

References:

1. Remus W, O'Connor M. Neural networks for time-series forecasting. In: Principles of forecasting. Springer; 2001. p. 245–256.
2. Li Q, Meng Q, Cai J, Yoshino H, Mochida A. Predicting hourly cooling load in the building: A comparison of support vector machine and different artificial neural networks. Energy Convers Manag. 2009;50(1):90–6.
3. Etheridge HL, Sriram RS, Hsu HYK. A comparison of selected artificial neural networks that help auditors evaluate client financial viability. Decis Sci. 2000;31(2):531–50.
4. Celikoglu HB. Application of radial basis function and generalized regression neural networks in non-linear utility function specification for travel mode choice modelling. Math Comput Model. 2006;44(7–8):640–58.
5. Harikumar R, Vijayakumar T. A comparison of elman and radial basis function (rbf) neural networks in optimization of fuzzy outputs for epilepsy risk levels classification from eeg signals. Int J Soft Comput Eng. 2013;2(6):295–303.
6. Li G, Shi J. On comparing three artificial neural networks for wind speed forecasting. Appl Energy. 2010;87(7):2313–20.
7. Zhao X, Han M, Ding L, Calin AC. Forecasting carbon dioxide emissions based on a hybrid of mixed data sampling regression model and back propagation neural network in the USA. Environ Sci Pollut Res. 2018;25(3):2899–910.
8. Zhang GP. Time series forecasting using a hybrid ARIMA and neural network model. Neurocomputing. 2003 Jan 1;50:159-75..
9. Xiong T, Li C, Bao Y. Seasonal forecasting of agricultural commodity price using a hybrid STL and ELM method: Evidence from the vegetable market in China. Neurocomputing. 2018;275:2831–44.
10. Zhang D, Martinez N, Lindholm G, Ratnaweera H. Manage sewer in-line storage control using hydraulic model and recurrent neural network. Water resources management. 2018 Apr;32(6):2079-98.
11. Ashour M, Al-Dahhan I, Hassan A. Forecasting by Using the Optimal Time Series Method. InInternational Conference on Human Interaction and Emerging Technologies Springer 2020 Apr 2: 148-154.
12. Abbas RA. Improving time series forecast errors by Using Recurrent Neural Networks. Proceedings of

the 7th International Conference on Software and Computer Applications. NY united states 2018: 229–232.

13. Abdul M, Ashour H, Jamal A, Alayham R, Helmi A. Effectiveness of Artificial Neural Networks in Solving Financial Time Series. Int J Eng Technol. 2018;7(4.11):99-105.

14. Tang Z, De Almeida C, Fishwick PA. Time series forecasting using neural networks vs. Box-Jenkins methodology. Simulation. 1991 Nov;57(5):303-10.

15. Krishna B, YR Satyaji R, PC N. Time series

modeling of river flow using wavelet neural networks. J Water Resour Prot. 2011 Jan 29;2011.

16. Ashour MA, Al-Dahhan IA. Turkish lira Exchange rate forecasting using time series models. Int e-j adv Soc Sci. 2020;6(16):294-300.

17. Tawfiq LNM. Stability of Back Propagation Training Algorithm for Neural Networks. Baghdad Sci J. 2012;9(4):713-719.

18. Tawfiq LN. On Training Of Feed Forward Neural Networks. Baghdad Sci J. 2007;4(1):158-164.

نماذج الشبكات العصبية الاصطناعية المثلى لسلاسل الزمنية

مروان عبد الحميد عاشور

قسم الإحصاء، كلية الإدارة والاقتصاد، جامعة بغداد، بغداد، العراق.

الخلاصة:

تعد الشبكات العصبية الاصطناعية أداة قوية وفعالة في تطبيقات السلاسل الزمنية. ان الهدف الأول في هذا البحث هو تشخيص أفضل واكفئ نماذج الشبكات العصبية الاصطناعي (الانتشار العكسي، دالة الأساس الشعاعي، الشبكة التكرارية) غفي حل السلاسل الزمنية الخطية وغير الخطية. اما الهدف الثاني هو التخلص من مشاكل الامثلية المحلية التي تعد من اهم مشاكل تقدير النماذج غير الخطية، واختبار حصانة نماذج الشبكات العصبية الاصطناعية. من اجل تشخيص افضل او امثل نماذج الشبكات العصبية في هذا البحث، استعمل مهارة التنبؤ لقياس كفاءة أداء نماذج الشبكات العصبية الاصطناعية. كما تم استعمال جذر معدل مربع الخطأ والمتوسط المطلق للخطأ النسبي لقياس دقة خطأ الطرائق المعتمدة. ان اهم ما تم التوصل اليه من خلال هذا البحث هو الشبكة العصبية الاصطناعية الأمثل كانت شبكة الانتشار العكسي والشبكة التكرارية لحل السلاسل الزمنية سواء كانت الخطية او غير الخطية. اثبت النتائج عدم كفاءة ودفة شبكة دالة الأساس الشعاعي لمعالجة السلاسل الزمنية غير الخطية، وكفاءتها في السلاسل الزمنية الخطية او شبه الخطية فقط، وعدم مقدرتها على التخلص من مشاكل الامثلية المحلية. النتائج المقدمة في هذا البحث تحسن من طرائق الحديثة لمعالجة التنبؤ بالسلاسل الزمنية.

الكلمات المفتاحية: كفاءة الشبكات العصبية، مهارة التنبؤ، الامثلية المحلية، الامثلية، السلسلة الزمنية.