Periodicals of Engineering and Natural Sciences ISSN 2303-4521
Vol. 8, No. 1, March 2020, pp.125-132

Curve fitting predication with artificial neural networks: A comparative analysis

Rabab Alayham Abbas[1], Marwan Abdul Hameed Ashour[2], Arshad Jamal[1], Sim Liew Fong[1]
[1] Faculty of Information Sciences & Engineering, Management & Science University
[2] College of Administration & Economics, University of Baghdad

ABSTRACT

Artificial neural networks (ANN) is considered one of the most efficient methods in processing Big Data, they have a great potential in economics and engineering applications. The aims of this paper is to discuss the best method for forecasting time series by comparing the results of ANN and Box and Jenkins methods (BJ) or ARMA models. As well as finding the best curve fitting and forecasting for linear or semi linear model. In this paper uses 3 error indicators to measure the efficiency of forecasting for the forecasting performance. The most important conclusion of this paper Proved that artificial neural networks are more effective than Box-Jenkins method or ARMA models in solving time series. The results also proved that artificial neural networks are significantly improving errors in the results and this is the ambition of all researchers.

Keywords: Big Data, ARMA model, Time Series, Artificial Neural Networks, Back propagation

Corresponding Author:

Marwan Abdul Hameed Ashour
Assistant Professor-Statistics Department-College of Administration & Economics
University of Baghdad
Iraq
+9647901571255
E-mail: dr_Marwan2012@yahoo.com

1. Introduction

Featuring the information revolution in the past few decades stands for a great advancement achieved in the field of data accusation and data processing, thus resulting in a new phenomenon termed as Big Data. With massive amounts of data being collected every second from various sources, those fitting under one if not more of the characteristics of big data, manly Volume, Variety, Velocity and Veracity. Therefore, emerges the necessity to find means to process and analyze such considerable data sets. Artificial Neural Networks (ANN) has received great attention from researchers and scholars as a method to handle such data sets, because of their ability to identify patterns in the data, thus forecast future data instances without the need for a predefined model.

Although, the work on synthesizing a model for data management from the biological nervous systems has a longer history, but the advancement in processing power abilities of hardware and distributed processing was the main reason it became feasible and practicable solution. Hence, it opens a wide horizon for research in the field of ANN to identify its quality and efficiency in data processing and analysis [1].

Artificial neural networks have the ability to solve any model, whether linear, semi-linear or non-linear, because of the property of self-learning. It is very useful to solve time series without the need to diagnose the model that is required by traditional models. In this paper, a free model has been proposed by the Back Propagation Network, which can solve any time series whether linear, semi-linear or nonlinear.

Many researchers has established that ANNs are one of the most accurate and widely used forecasting models that have enjoyed fruitful applications in forecasting, social, economic, engineering, foreign exchange, stock problems, etc. In general, neural networks consist of three layers: input, hidden and output. Each layer has a nodes' set, this layer is connected to weights representing neural dendrites, and the computational algorithms are followed in the contract of the logistics function. [2-4]. On the other hand, the Box and Jenkins methods (BJ) or ARMA models are good classical methods, to be compared with ANN, which are very powerful. One of the Defects of the Box and Jenkins method (BJ) is its reliance on past observations and the misdiagnosis of the model, while artificial neural networks have the property of self-learning and training with any model or time series [5] [6] [7] [8]. The approach of using the different types of ANN has been applied in various filed of science as in the work of numerous researchers [9], [10], [11], [12]. Moreover the selection of the ANN as a method for improving forecasting has been adopted by a significant number of researchers [13-15]. It was the application of a well-known real data set was the energy data for Iraq crude oil production from 1998 to 2015. To reach the best curve-fitting model or series by using traditional and modern methods, where oil in Iraq and the Middle East is the mainstay of the national economy and the main and sole source of Budget income for the country. The rest of the paper covers the methodology, practical application, results' discussion and conclusion.

2. Methodology

In this section, the methodology phases to conduct this research are introduced. The first phase is to provide a review of the relevant theoretical base for the research, thus covering the time series and methods used to handle them, the proposed model and the classical model, in order to set the basic theoretical background for the application of those methods. While the second phase covers practical application of the models, it also includes the statistical measurements used to compare the efficiency and suitability of both models. As well as an introduction to ANN and the manner in which it can be used in such type of series. It also includes a comparison between the methods presented. The third phase is to compare the results from the models and elaborate on the impact of the findings; Figure 1 shows the methodology phases of this research.

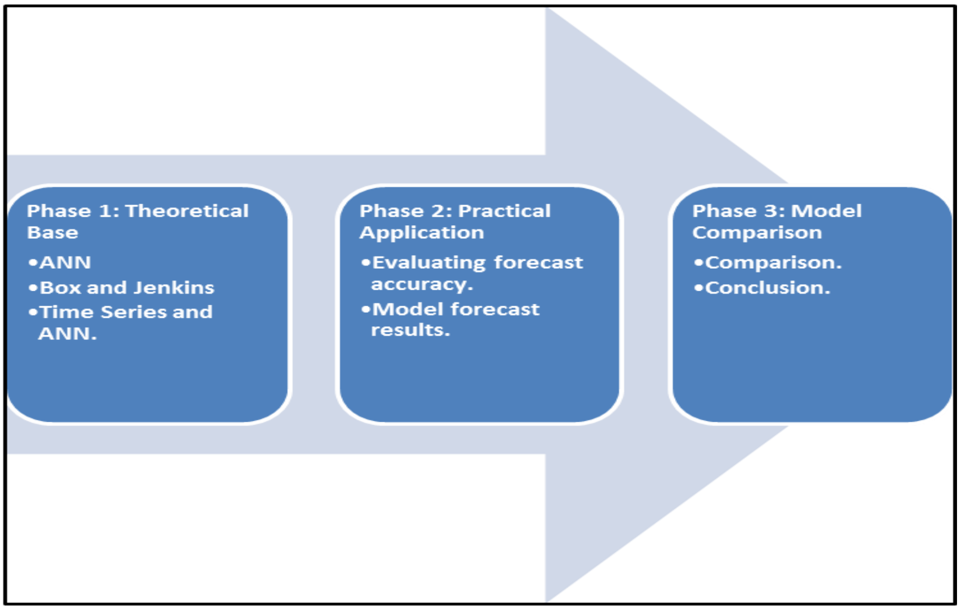

Figure 1. Research methodology

3. Theoretical base

3.1. Artificial neural networks

Artificial neural networks are a nonlinear model, derived from the biological neural systems. The name of ANN refers to the network being a combination of an interconnected and communicating unit; those units were originally inspired by the study of the nervous systems in living creatures. The manner in which the ANN work is based on the idea of reaching a model through simulating the data, in order to see that model for analysis, classification, estimation and other data processes, without the need for developing a model for the data. Therefore, ANN succeeded in gaining the attention of a large number of scientists and researchers. This increasing attention and focus on ANN is due to its high flexibility in the data model learning process as well as storing and distribution of data, in comparison with other mathematical methods used in the process. The back propagation network is one of the best types of neural networks and the most prevalent. The basic steps of the Back Propagation Methodology (BPM) is to calculate the error in the output layer, then use it to update the weightage for the hidden-output layer. Similarly, the error for the hidden layer is computed to update weightage for the input-hidden layer. After that, the output for the network using the new wattage is calculated and the same process steps of calculating error and updating weightage were repeated until minimal errors for the network is reached. The methodology of Back Propagation networks can be described using the following steps [3]–[5] :

- Generating initial weighted random– from one of the statistical distributions.
- Feed-forward computation

i. Compute values of the hidden nodes

$$h_j = 2/(1 + \exp -(\sum_i (x_i.w_{ij})) - 1 \quad (1)$$

ii. Compute values of the output nodes

$$y_k = 2/(1 + \exp -(\sum_j (h_j.w_{jk})) - 1 \quad (2)$$

- Back-forward computation

Compute the error value for the output node

$$E_k = t_k(\text{target}) - y_k(output) \quad (3)$$

- Compute output nodes

$$\delta_K = E_K.y_K(1 - y_K) \quad (4)$$

- Compute hidden nodes

$$\delta_j = h_j.h_j(1 - h_j) \quad (5)$$

- Update weights for each input-to-hidden layer Wright W_{ij}, compute the size change

$$\Delta w_{ij} = \eta\delta_j x_j + \alpha\Delta w_{ij} \quad (6)$$

Compute the new weight

$$w_{ij} = w_{ij} + \Delta w_{ij} \quad (7)$$

For each hidden-to output layer weight W_{jk}, compute the size of the weight change

$$\Delta w_{jk} = \eta\delta_k h_k + \alpha\Delta w_{jk} \quad (8)$$

Compute new weight

$$w_{jk} = w_{jk} + \Delta w_{jk} \quad (9)$$

- Test network generalization and Stop condition.

3.2. Box and Jenkins methods *(BJ)*

Time series defined as an observed values sequence for a specific time related random phenomenon. Some statistical criteria are used to describe the quality of the time series and thus to facilitate their modeling. The autocorrelation function (ACF) indicates the relationship or correlation of adjacent string values and the autocorrelation coefficient between -1 and 1. The partial autocorrelation function (PACF) measures the partial effect of adding the delayed values of a variable, and PACF coefficients can be obtained from the auto-regression equation of the subject. The best predictor of the value of X_t in time to + L follows the mixed model ARMA (p, q) [7]–[9].

$$x_t = \mu + \varphi_1 x_{t-1} \dots + \varphi_p x_{t-p} - a_t - \theta_1 a_{t-1} \dots - \theta_q a_{t-q} \qquad (10)$$

3.3. Time series and artificial neural networks

Neural Networks are gaining increasing importance in processing and analyzing time series, due to its ability to self-learning and training. Most researchers use theorems based on shifting or lag the time series through one degree or more, to select the input time series artificial neural network that is solving the time series, The neural networks inputs are determined by the following formula [2], [4].

$$x_t = w_0 + \sum_{j=1}^{q} w_j . g\left(w_{0j} + \sum_{i=1}^{p} w_{ij} . x_{t-i} \right) + \varepsilon_t \qquad (11)$$

4. Practical application

In this part the two methods, classic and suggested are applied to the selected data set, later the results of the forecast for both the models are compared to decide the most accurate method to forecast the results, the accuracy of the forecast is based on the following criteria

4.1. Evaluating forecast accuracy

For comparison purposes, the results of the methods adopted, can adopt the following statistical criteria[4], [7], [8].

- Mean absolute percentage (MAPE)

$$\text{MAPE}=\frac{\sum\left|\frac{et}{xt}\right|}{n} * 100 \qquad (12)$$

- Root mean squared error (RMSE)

$$\text{RMSE}=\sqrt[2]{\frac{\sum et}{n}} \qquad (13)$$

Where: x_t is actual value, $\hat{x}_t$ is forecast value, n: is Sample size, a_t is error and.

$a_t = x_t - \hat{x}_t$,

- Coefficient of determination (R^2)

$$R^2 = \frac{\sum X_t - \bar{X}}{\sum \hat{X}_t - \bar{X}} \qquad (14)$$

4.2. Application data set

In this section, used well-known real data sets for the energy data on Iraq crude oil production from 1998 to 2015 is shown in Figure 2 to reach the best curve-fitting model or series by using traditional and modern methods, where oil in Iraq and the Middle East is the mainstay of the national economy. In addition, the main and sole source of budget income for the country. In addition, the MATLAB (Version 8.2) high-level programming language and SPSS programming (V.21) are used for this implementation.

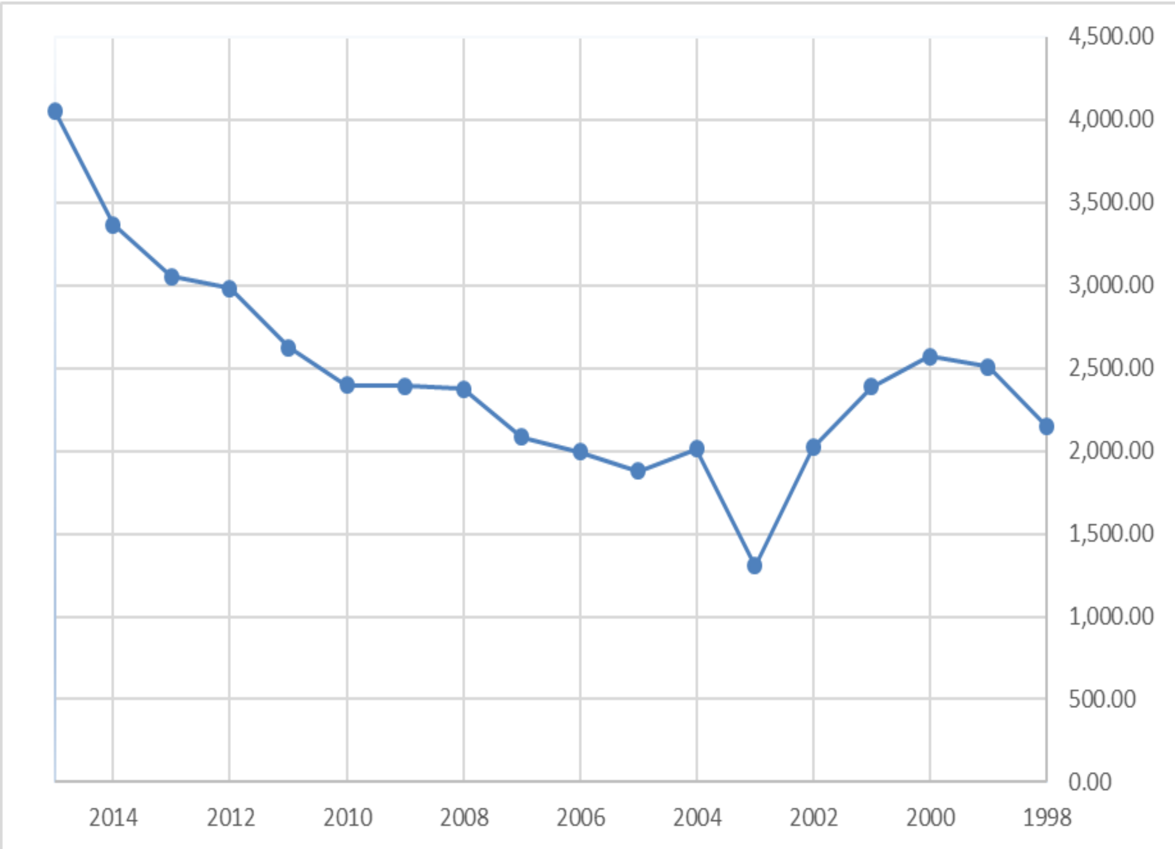

Figure 2. Iraq oil production series (barrels per specified year) as stated by United States Energy Information Administration

5. Results and discussion

It is shown in Figure 3 that the ACF is decreasing exponentially and the PACF is cut off at the first leg, so the best ARIMA model for Iraq oil production based on ACF and PACF is an ARIMA (1,0,0) and the result is shown in Table 1 by SPSS programming (V.21). Then, the mathematical model is:

$$x_t = 0.898x_{t-1} + \varepsilon \qquad (15)$$

Moreover, Table 1 clearly shows that all parameters are significant.

Table1. ARIMA model parameters

	Estimation Parameters	Sig.
ϕ	0.898	0.000

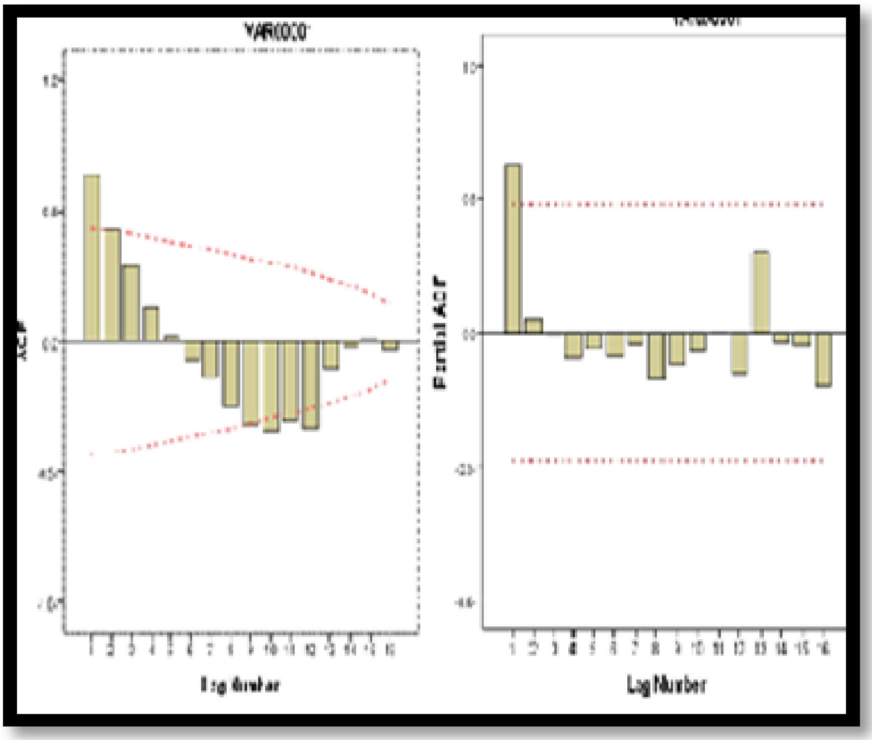

Figure 3. ACF & PACF for time series

The behavior of ACF & PACF for residual is shown in Figure 3, the error behavior is random. In addition, curve fitting for ARMA models is shown in Figure 4.

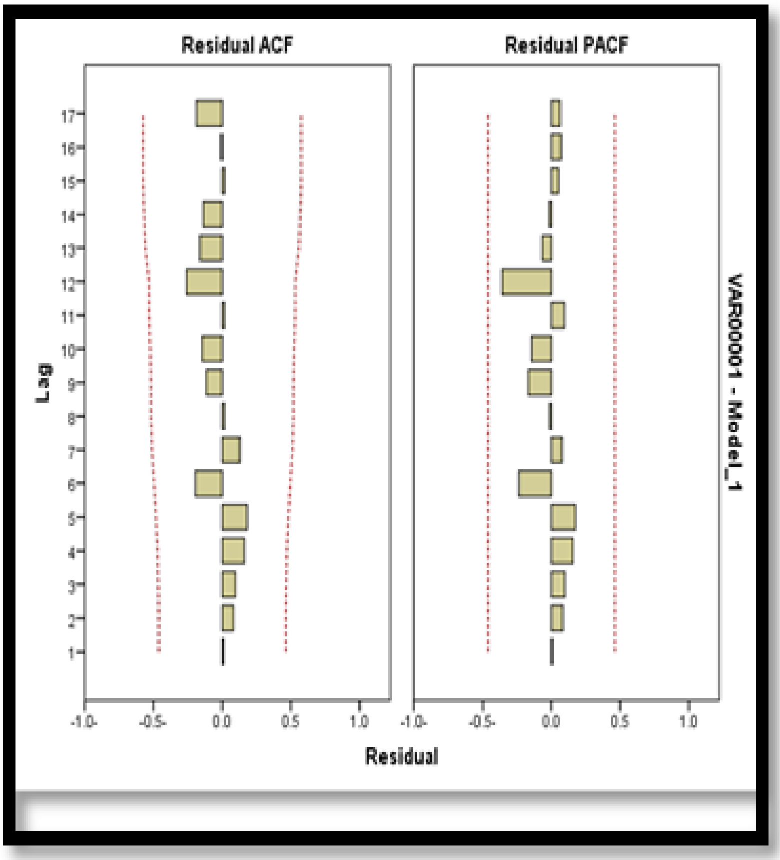

Figure 4. ACF & PACF for residual

The Back Propagation Network architecture is as follows using MATLAB simulator.

- Input layer- based to the equation (11) and since the ARIMA model are (1,0,0), the number of nodes in this layer is one node, with two degrees' time series shift, represented by the variables (X_{t-1}) as follows:

$$x_t = w_0 + \sum_{j=1}^{q} w_j . g\left(w_{0j} + \sum_{i=1}^{p} w_{ij}.x_{t-1}\right) + \varepsilon_t \quad (16)$$

- Network Training: In the training phase, a logistic function and a sigmoid function, and the maximum number of iterations is 1500, unless the one of the training condition was met. The best performance of the network, and best training was decided based on finding the value of Regression (R).
- Hidden Layer: In this layer, the optimal number of nodes is 15 nodes, with one layer (after many trials).
- Output Layer: This layer consists of one node only, represented by the variable X_t. The performers results for the BP network for the time series under study are clear from the results of the error criteria to compare the two methods used is shown in Table 2. While histogram of the MAPE for two methods is shown in Figure 5.

Table2. Results of ANN

Method / Carrion	RMSE	MAPE	R^2
ARMA (B&J)	398.971	12.979	0.612
ANN	2.4091	0.3280	0.9989

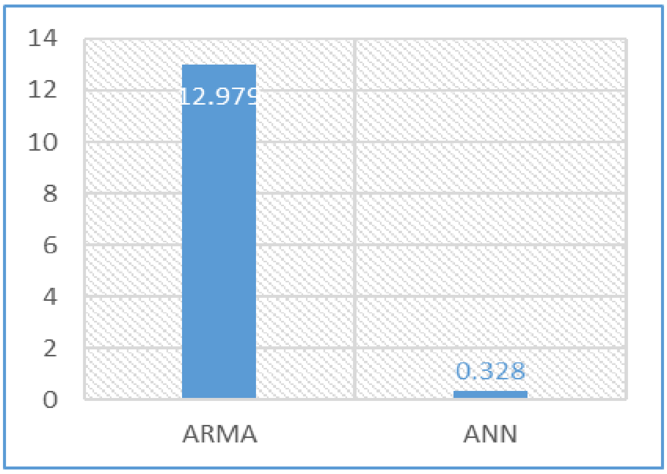

Figure 5. Histogram of MAPE for two methods

6. Conclusion

For evaluation purpose, three typical indicators RMSE, MAPE and R^2 had used to compare predicted and actual values to validate each forecasting system. The results of the methods were used in this paper are is shown in Table 2 and Figure 8. The best method or curve fitting for time series of Iraq oil production that the method that give minimized error possible and the largest of the coefficient of determination, was artificial

neural networks and is the best Box and Jenkins method. Results showed also we had been seen that the neural networks using a traditional model (Box and Jenkins method or ARMA models) were good.
It is also clear that the difference in the error results was largest between the modern model and the traditional model because Artificial neural networks have a property self-learning and training with any model which gives efficient results, while the traditional method that uses past observations for the series. The results also proved that the difference between the two methods was significantly in error criteria. Artificial neural networks improve the forecast results error significantly and this is the ambition of all researchers.
The results concur with other researches that are using the ANNs and comparing their results with the classical model as in the work presented in researches [13-15].

References

[1] H. Demuth, “Neural Network Toolbox,” Networks, vol. 24, no. 1, pp. 1–8, 2002.
[2] KHASHEI, Mehdi; BIJARI, Mehdi. “An artificial neural network (p, d, q) model for time series forecasting”. Expert Systems with applications, vol. 37, no. 1, pp 479-489, 2010.
[3] Samarasinghe, Sandhya. Neural networks for applied sciences and engineering from fundamentals to complex pattern recognition. CRC Press, 2016.
[4] P. G. Zhang, “Time series forecasting using a hybrid ARIMA and neural network models,” Neurocomputing, vol. 50, no. February, pp. 159–175, 2003.
[5] Chatfield, Chris. “The analysis of time series an introduction”. CRC press, 2016.
[6] BROCKWELL, Peter J.; DAVIS, Richard A. “Introduction to time series and forecasting”. Springer, 2016.
[7] CHATFIELD, Chris. The analysis of time series an introduction. CRC press, 2016.
[8] S. K. Safi, “Using ANNs and ARIMA Models to Make Accurate Forecasts for Palestinian Official Statistics Based on Simulation and Empirical Applications,” vol. 25, no. 2, pp. 106–116, 2017.
[9] Laptev, Nikolay, et al. "Time-series extreme event forecasting with neural networks at uber." Int. Conf. on Machine Learning. 2017.
[10] O. Erdogan and A. Goksu, "Forecasting Euro and Turkish Lira Exchange Rates with Artificial Neural Networks (ANN)," *International Journal of Academic Research in Accounting, Finance and Management Sciences*, vol. 4, pp. 307-316, 2014.
[11] F. Y. H. Ahmed and S. M. Shamsuddin, "The Effect of Weights Initialization on Osteo Arthritis Classification," in 2010 Fourth Asia International Conference on Mathematical/Analytical Modelling and Computer Simulation, pp. 87-92, 2010.
[12] A. Atul and L. Suganthi, "Forecasting of Electricity Demand by Hybrid ANN-PSO Models," *International Journal of Energy Optimization and Engineering (IJEOE)*, vol. 6, pp. 66-83, 2017.
[13] M. A. H. Ashour and R. A. Abbas, "Improving Time Series' Forecast Errors by Using Recurrent Neural Networks," presented at the Proceedings of the 2018 7th International Conference on Software and Computer Applications, Kuantan, Malaysia, 2018.
[14] M. A. H. Ashour, A. Jamal, and R. A. A. Helmi, "Effectiveness of Artificial Neural Networks in Solving Financial Time Series," International Journal of Engineering & Technology, vol. 7, pp. 99-105, 2018.
[15] C. N. Babu and B. E. Reddy, "A moving-average filter based hybrid ARIMA–ANN model for forecasting time series data," Applied Soft Computing, vol. 23, pp. 27-38, 2014.