

Comparative analysis of short-term demand predicting models using ARIMA and deep learning

Halima Bousqaoui¹, Ilham Slimani², Said Achchab³

^{1,3}National Higher School for Computer Science and System Analysis, Mohammed V University, Morocco

²SmartICT Laboratory, ENSAO, Mohammed First University, Oujda, Morocco

Article Info

Article history:

Received Sep 9, 2020

Revised Dec 11, 2020

Accepted Jan 13, 2021

Keywords:

ARIMA

Artificial neural network

Convolutional neural network

LSTM

Machine learning

Supply chain management

ABSTRACT

The forecasting consists of taking historical data as inputs then using them to predict future observations, thus determining future trends. Demand prediction is a crucial component in the supply chain's process that allows each member to enhance its performance and its profit. Nevertheless, because of demand uncertainty supply chains usually suffer from many problems such as the bullwhip effect. As a solution to those logistics issues, this paper presents a comparative analysis of four time series demand forecasting models; namely, the autoregressive integrated moving Average (ARIMA) a statistical model, the multi-layer perceptron (MLP) a feedforward neural network, the long short-term memory model (LSTM) a recurrent neural network and the convolutional neural network (CNN or ConvNet) a deep learning model. The experimentations are carried out using a real-life dataset provided by a supermarket in Morocco. The results clearly show that the convolutional neural network gives slightly better forecasting results than the Long short-term memory network.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



Corresponding Author:

Halima Bousqaoui

National Higher School for Computer Science and System Analysis

Mohammed V University

Ben Abdallah Regragui, Madinat Al Irfane, BP 713, Agdal Rabat, Morocco

Email: halima.bousqaoui@gmail.com

1. INTRODUCTION

A supply chain is an entire network composed of distributed organizations in interaction (suppliers, retailers, distributors, and clients) involved in the process of purchasing, inventory management, production, distribution and delivery of a product or a service from raw materials to a final customer [1]. Supply chain management covers information flows, physical distribution as well as financial transactions. In other words, all the supply chain members share a common objective that is improving their businesses and profits while creating a value represented by the product or service delivered to the customer [2].

The problem consists of the demand forecasting which is a crucial component in supply chain's process. Indeed, prediction consists of taking models from historical data in order to use them to predict future observations. Nevertheless, the uncertain character of the customer's demand causes difficulties for decision makers who must react in an efficient and fast manner. Besides, with the advancement of technology, supply chains become more complex. So, in order to keep up with this vital evolution and for an effective supply chain, companies must be able to plan not only the present but also the future sales or demand. Certainly, sales forecasting is efficient for supply chains' planning activities including purchasing, inventory, production and distribution. It helps managers to make optimal decisions.

As background in the literature, autoregressive integrated moving average (ARIMA), is a linear time series forecasting method that has been widely used in many areas to model different problems. Some of the applications of ARIMA are as: Fanoodi *et al.* [3] uses ARIMA to reduce demand uncertainty in a blood platelet supply chain. Ji *et al.* [4] uses a hybrid ARIMA and deep neural network model to forecast future prices of carbon. Matyjaszek *et al.* [5] implements an ARIMA model in order to forecast coking coal prices in the mining industry. Ohyver and Pudjihastuti [6] uses the ARIMA model to forecast rice price. These are only a few examples of the application of ARIMA that remains widely employed in multiple domains to make accurate forecasts. ARIMA is also used for demand forecasting. For instance, Oliveira *et al.* [7] proposes a double seasonal ARIMA model to make accurate short-term water demand forecasts. Al-Musaylh *et al.* [8] applies ARIMA, MARS, and SVR in short term electricity demand forecasting. Anggraeni *et al.* [9] compares between ARIMA and ARIMAX results for kids clothes demand forecasting. Amini *et al.* [10] applies ARIMA for an electric vehicle charging demand time series forecasting. And Wang *et al.* [11] implements a modification approach in seasonal ARIMA for electricity demand forecasting.

Neural networks on the other hand, are used for nonlinear time series modeling. Some applications in demand forecasting include: In the paper [12], neural networks are applied for demand forecasting of an engine oil. Sharma *et al.* [13] applies neural networks to predict energy demand, carbon dioxide emissions and wind generation. Matino *et al.* [14] implements neural network-based models for a blast furnace gas production and demand prediction problem. And Silva *et al.* [15] uses denoised neural networks to forecast tourism demand. In [16], Convolutional neural network is applied to tackle the problem of predicting short-term supply-demand gap of ride-sourcing services, where three hexagon-based convolutional neural networks (H-CNN) are proposed. Ke *et al.* [17], propose a hybrid approach or the fusion convolutional long short-term memory network, where multiple convolutional long short-term memory (LSTM) layers, standard LSTM layers, and convolutional layers are fused aiming to forecast passenger demand under an on-demand ride service platform. Amarasinghe *et al.* [18] compares convolutional neural networks to perform energy load forecasting based on historical loads of a single residential customer to long short-term memories (LSTM) sequence-to-sequence (LSTM S2S), factored restricted Boltzmann machines (FCRBM), artificial neural networks (ANN) and support vector machines (SVM). The obtained results demonstrate that the CNN outperformed SVR while producing comparable results to the neural network and deep learning models.

The present paper tackles the proposed solution which is using time series forecasting models and deep learning models for demand prediction. In addition, the convolutional neural network has been rarely used in the literature. Using it in comparison with other models consists of the main contribution of the paper. The novelty presented in this paper is the performance evaluation and comparative study using a real dataset provided by a supermarket in Morocco. The predictive performances of four prediction methods are presented. Namely, the autoregressive integrated moving average (ARIMA) as statistical model, the multi-layer perceptron (MLP) a feedforward neural network [19, 20], the convolutional neural network (CNN or ConvNet) a deep learning model and the long short term memory model (LSTM) a recurrent neural network. These models were chosen in order to compare between the different types of neural networks. Mainly the feedforward, the convolutional and the recurrent networks.

The remainder of this paper is organized as follows: The second part presents an overview of the proposed prediction methods based on time series methods. Then, the third section is dedicated to explain the adopted methodology and illustrate the obtained forecasting results in a comparative analysis based on numerical experimentations using a real dataset of a local supermarket. Finally, the concluding remarks and the future work are presented.

2. PROPOSED PREDICTION METHODS

Time series can be defined as a series of data points recorded and analyzed in a time order, it is a sequence taken at equally spaced time periods [21]. The only independent variable in time series methods is time. The analysis of time series is the process of extracting meaningful information and statistics from the data. The forecasting of time series is the process of estimating the future values of the data based on previously observed data plus some other features [22]. Contrary to supervised models forecasting where data is predicted using multiple features, in time series forecasting, a variable is predicted using past observations of the same variable. The Figure 1 shows different methods used in time series forecasting: Forecasting techniques are categorized into two families:

- Parametric methods based on mathematical tools and statistical analysis using historical data such as linear and nonlinear regression, autoregressive integrated moving average. Nevertheless, those methods are complicated to use.

- Non-parametric methods based on machine learning with the ability of learning and approaching any nonlinear function. Those techniques are mostly based on the use of artificial intelligence such as artificial neural networks that offer flexible parameters during the learning and implementation phases.

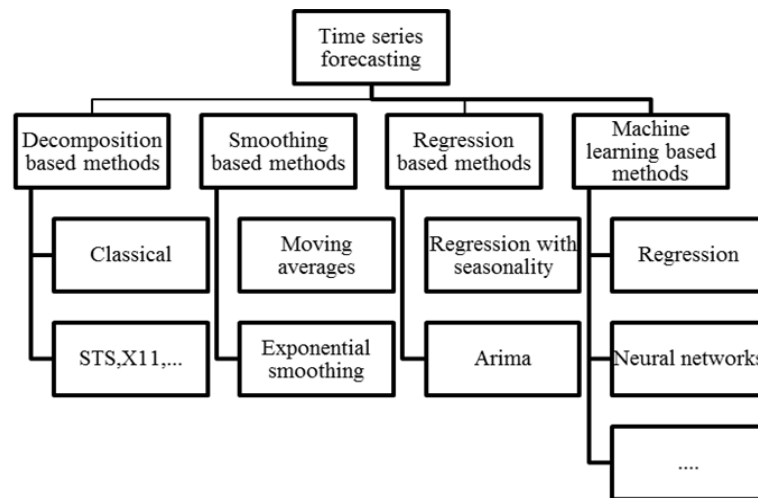


Figure 1. Time series forecasting methods [22]

2.1. Autoregressive integrated moving average (ARIMA)

The ARIMA methodology was originally developed by Box and Jenkins in the 1970s. A model is denoted by ARIMA (p,d,q) and is composed of:

- An Autoregressive model of order p (number of lags).
- Integrated refers to the step of removing non stationarity by differencing the data. d is the degree of differencing.
- A Moving average model of order q.

2.1.1. Autoregressive model

The autoregressive model AR(p) of order p states that the output variable depends on its previous values with some lag plus a random term. The autoregressive model AR(p) is defined as:

$$y_t = A + \alpha_1 y_{t-1} + \alpha_2 y_{t-2} + \dots + \alpha_p y_{t-p}$$

2.1.2. Differencing

A time series is called non stationary when the values of the data are dependent on time. A non-stationary time series can be changed into stationary by differencing the time series. Differencing refers to subtracting some past values of the data a number of times. d the degree of differencing refers to that number of times. Generally, a series might need first-differencing multiple d times to attain stationarity [23]. A differencing of degree 1 is as:

$$y'_t = y_t - y_{t-1}$$

2.1.3. Moving average model

To deal with non-stationarity, we characterize the time series as the sum of a non-constant mean value plus a random error variable [23]:

$$x_t = \mu_t + \varepsilon_t$$

In smoothing methods, a variable is a function of some past observations, which means that the future value of the time series is the weighted average of some past observations. As such, the moving average MA(q) of order q can be written as the weighted average of the past q errors.

$$y_t = B + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

2.1.4. Autoregressive integrated moving average

ARIMA is a mixed model that combines both the differenced autoregressive and moving average models. The final form of a time series model, which depends on its own p past values and on the q past values of white noise error terms [21], is as:

$$y_t = I + \alpha_1 y'_{t-1} + \alpha_2 y'_{t-2} + \dots + \alpha_p y'_{t-p} + \varepsilon_t + \theta_1 \varepsilon_{t-1} + \theta_2 \varepsilon_{t-2} + \dots + \theta_q \varepsilon_{t-q}$$

with y_t differenced to a degree d . This represents the ARIMA(p,d,q) model.

2.2. Machine learning based methods

The artificial neural networks are popular machine learning algorithms that simulate how the human brain behaves and learns. Neural refers to neurons, they are cells contained in the human nervous system. The neurons are connected to one another with the use of dendrites and axons, while synapses are the regions connecting between these axons [24].

Artificial neural networks (ANN) are an artificial intelligence method inspired from the functioning of the biological neural networks characterized by a circuit of interconnected neurons organized as multiple interconnected layers: an input layer, one or more hidden layers and an output layer. ANNs are very powerful in the way that they can identify complex linear/non-linear relationships between inputs and outputs. Neural networks are used in various fields from classification, to forecasting, to approximation, to diagnosis, to image processing or even recognition.

There are two types of neural networks:

- Feedforward neural network, or also called non-recurrent neural network where information flows only in one direction going from the input to the output layer all the way through the hidden layer or layers. However, they have no ability to maintain historical inputs as they only consider the current inputs.
- Recurrent neural network, where information cycles in a loop in both directions, from the input layer to the output layer and the other way around. Some networks allow the persistence of information by taking into account the current input as well as inputs received previously.

2.2.1. Multi-layer perceptron (MLP)

This specific architecture of multi-layer neural networks is called a feed-forward neural network, because multiple “successive layers feed into one another in the forward direction from input to output” [24]. Therefore, the multi-layer perceptron network generates an output from some given inputs. They consist of multiple layers of nodes or neurons (at least three). One input layer, some hidden layers, and one output layer. Every node except the input ones are called neurons. Nonlinear activation functions are used to model the non-linearity of a given problem. The computation happens in the hidden layers.

Firstly, the first layer sends input data to the hidden nodes in the hidden layer. These nodes combine the data with a set of coefficients or weights that either amplify or minimize the input, then the resulting products are added. Finally, an activation function is applied to the sum. The activation function determines how much and whether the signal progresses through the network to affect the final result.

2.2.2. Long short-term memory (LSTM)

The LSTM network was introduced by Hochreiter and Schmidhuber in 1997 [25]. It is a recurrent neural network, specifically a gated network used in deep learning. Meaning it has feedback connections unlike the feedforward neural networks. It can be used in classification and forecasting of a time series and it is especially effective in solving sequential problems like speech and handwritten recognition.

LSTMs are able to learn long-term dependencies; in other words, they excel at remembering information for long periods of time. A LSTM unit is composed of a cell (that has a self-loop), an input gate, an output gate and a forget gate. The cell remembers values over intervals of time while the three gates control the flow of information into and out of it. An example of a cell unit is shown in the Figure 2.

$$g = \tanh(U_g x_t + V_g h_{t-1} + b_g). \quad (1)$$

$$i = \sigma(U_i x_t + V_i h_{t-1} + b_i). \quad (2)$$

$$f = \sigma(U_f x_t + V_f h_{t-1} + b_f). \quad (3)$$

$$S_t = S_{t-1} \otimes f + g \otimes i \quad (4)$$

$$o = \sigma(U_o x_t + V_o h_{t-1} + b_o). \quad (5)$$

$$h_t = \tanh(S_t) \otimes o \quad (6)$$

The cell receives as input the current input x_t and the previous cell output h_{t-1} . It outputs h_t .

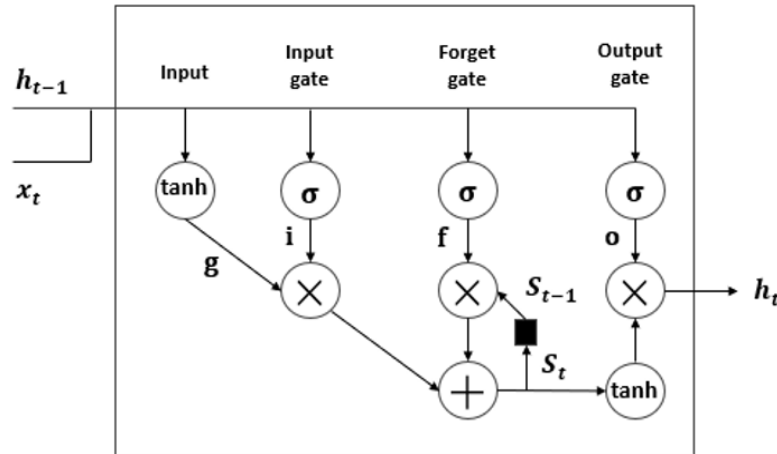


Figure 2. The structure of a typical memory cell of LSTM [26]

2.2.3. Convolutional neural networks (CNN or ConvNet)

A convolutional neural network is a specific type of artificial neural network usually used for cognitive tasks such as image recognition, image processing and natural language processing and time series data and more. CNN is the most popular deep learning algorithm model used for image processing, it is simpler to train and more effective than the traditional neural networks since it has the ability to capture the temporal and spatial dependencies in an image by applying the appropriate filters with an independence from human effort. Indeed, the layers in CNN architecture are arranged in order to cover the entire visual field that helps to avert the piecemeal image-processing problem in traditional neural networks.

Similar to the multilayer perceptron (MLP) structure of neural networks, the layers in CNN are categorized into three types: an input layer, an output layer and a hidden layer. There are two types of hidden layers [27]:

- Feature learning layers, performing three types of operations: convolution, pooling, and rectified linear unit (ReLU) on the input data;
- Classification layers, composed of fully connected layers and normalization layers.

3. EXPERIMENTATION RESULTS AND DISCUSSION

3.1. Autoregressive integrated moving average (ARIMA)

3.1.1. The proposed approach

The ARIMA (p,d,q) model is determined by choosing the three parameters p,d, and q. The plot of the data will determine the order of differencing d, the ACF plot will determine the lag q, and the PACF plot the lag p. Besides, the model is improved by using the grid search in order to choose the best p,d,q parameters. The models are evaluated with the RMSE, the Root Mean Square Error, that is the most frequently used metrics of prediction performance and calculated as the square root of the Mean Square Error, where the MSE is the arithmetic mean of the squares of the differences between forecasts and observations.

3.1.2. The ACF and PACF plots

In the Figure 3, demand is presented as a time series data having the day in the x-axis and the demand quantity on the y-axis. We notice that the time series might be stationary and so might not require differencing, but to be sure, at least a difference order of 1 will be applied. Next let us look at the autocorrelation function (ACF) and partial autocorrelation function (PACF) plots of the time series in Figure 4. The ACF plot is designed to show whether the elements of a time series are positively correlated,

negatively correlated, or zero correlated (independent of each other). In other words, it describes how well the present values in the series is related to its past values. A PACF is a summary of the relationship between observations in a time series with observations at prior time steps after excluding the contributions from the intermediate lags.

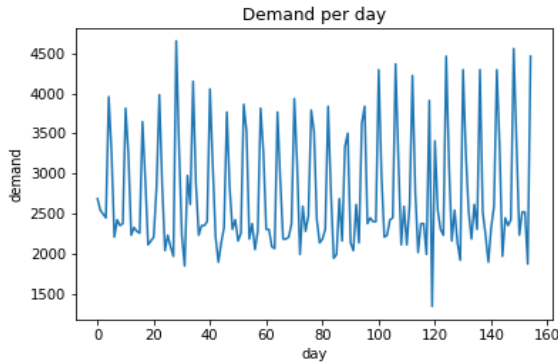


Figure 3. Demand per day plot

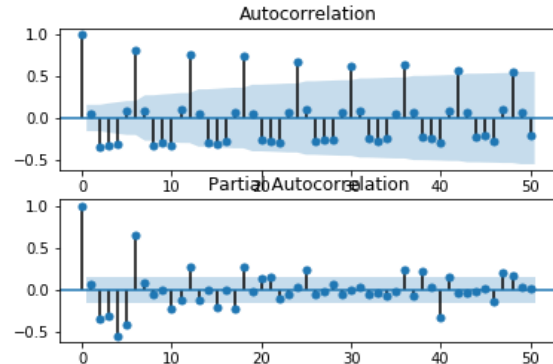


Figure 4. ACF and PACF plots for demand quantity series

We initially take the order of AR term to be equal to as many lags that crosses the significance limit in the PACF plot. This suggests that a good start for the autoregressive portion of the model is AR(2) that is an auto-regression model with two lag observation used as input. In the same way that we looked at the PACF plot for the quantity of the AR part of the model, we can look at the ACF plot for the number of MA terms. Only one lag is above the significance line. So let us fix the order as MA(1).

3.1.3. Fitting the ARIMA(2,1,1) model

Now that we have determined the values of p, d and q, we can fit the ARIMA model. Using the ARIMA(2,1,1) implementation in statsmodels package in Python we fit the model and we get the following results. The Figures 5 and 6 show that the residual errors are near zero mean with a uniform variance. Next, let us calculate the root mean squared error (RMSE) and plot the actual vs predicted values. Using sklearn and statsmodels packages of python we fit an ARIMA(2,1,1) model. This results in an RMSE of 823.945.

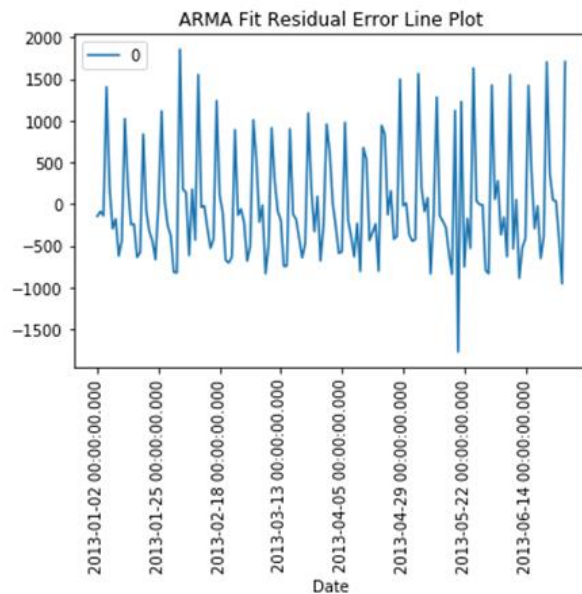


Figure 5. ARIMA fit residual error line plot

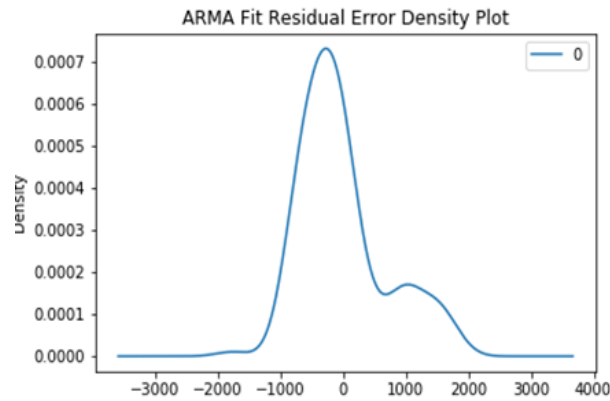


Figure 6. ARIMA fit residual error density plot

A line plot in Figure 7 is created showing the expected values (blue) compared to the forecast predictions (red). We can conclude that the model is in need of further tuning of the p , d , and q parameters. For a better fit of the model to the data, other parameters p , d , and q are tested using the grid searching parameters approach.

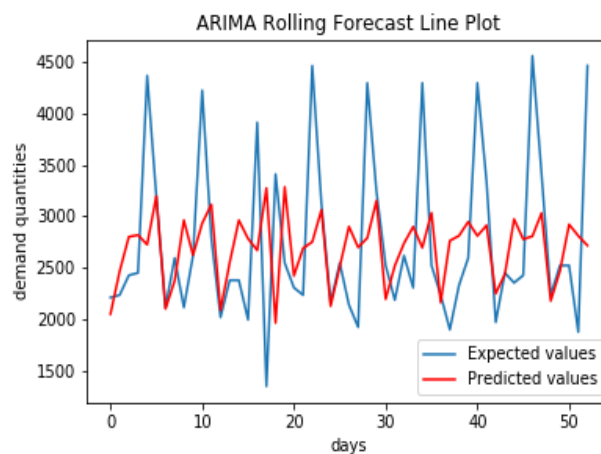


Figure 7. Predicted vs actual demand quantities

2.1.4. Grid searching parameters

The process of training and evaluating ARIMA models on different combinations of model hyperparameters can be automated. This is called a grid search or model tuning in machine learning. We split the data set in two: 66% for the training set and 34% for the test set. We chose the p values to range from 1 to 10, the d values from 0 to 3 and the q from 0 to 3. The best ARIMA model that yielded the best RMSE is ARIMA(6, 0, 0) with an RMSE of 485.690. This is shown in the Figure 8. We notice that the lower the p value the higher the RMSE.

3.2. Multi-layer perceptron (MLP)

The multi-layer perceptron neural network is used to forecast the demand quantity in a simple supply chain composed of one retailer and one supplier. The data belongs to a Moroccan supermarket dataset. Predicting the demand of a given day d involves the demand quantities of the three previous weeks ($d-7$, $d-14$ and $d-21$).

The neural network is composed of three hidden layers each containing 10 neurons. The neural network is trained using the adam version of the gradient descent, the mse loss function and the relu function as an activation function. This results in an RMSE of 0.140 in the test set after normalization and 464.261 before normalization.

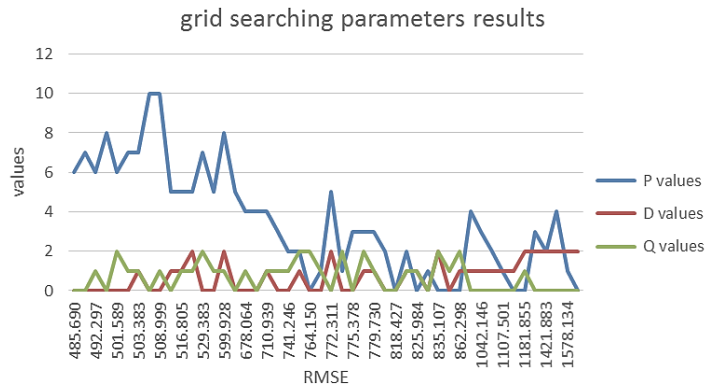


Figure 8. Grid searching parameters results

3.3. Long short-term memory (LSTM)

In a previous study [26], Long-Short Term Memory is used to forecast demand. In addition to the inputs proposed in the MLP model, the LSTM has a time dimension. It has a three-dimensional shape composed of the number of examples, the number of inputs, and the number of time steps [28].

In the proposed neural network, the demand of a given day “d” is predicted using demand quantities of the days: d-6, d-12, and d-18. Therefore, we model the problem as a three time-steps prediction on one input corresponding to the quantity of the day, and this for each output. This instead of considering that the data is composed of three inputs as we did in the MLP study.

The LSTM network is composed of 50 neurons in the hidden layer and 1 neuron in the output layer. The root mean squared error and the efficient Adam version of the stochastic gradient is used. This yielded an RMSE after normalization of 0.138 and an RMSE before normalization of 457.958.

3.4. Convolutional neural network (CNN or ConvNet)

The convolutional neural network (CNN or ConvNet) is a special type of neural network model used for working with two-dimensional and three-dimensional data like images, it can be applied with one-dimensional data like time series. The proposed convolutional neural network is composed of the following layers:

- A one-dimensional convolutional layer with 64 filters, a kernel size of 2 and an input shape of (3,1), the activation function consists of the relu function;
- A one-dimensional max pooling layer with a filter of size 2;
- A flatten layer;
- A fully connected layer composed of 50 neurons with relu being the activation function.

We train the model using the adam version of stochastic descent and the MSE loss function. The training after 200 epochs results in a RMSE of 0.138 after normalization, and 457.079 before normalization.

3.5. Discussion

In machine learning the prediction accuracy is used for discrete problems, meaning classification ones. Since we are dealing with a regression problem, we chose to evaluate the models using the root mean squared error RMSE, in addition to some other criteria. The RMSE measures how close the real values are to the predicted ones. Hence its use.

As shown in Table 1, the worst prediction results are obtained with ARIMA model which is understandable since it is a linear model. The neural networks proposed models, on the other hand, are nonlinear models that can approximate any function, hence the better results. CNN has shown the best RMSE, slightly better than the LSTM model. This might be thanks to its ability to do feature selection on its own.

Table 1. Results of the four approaches according to different criterias

	ARIMA	MLP	LSTM	CNN
RMSE before normalization	485.690	464.261	457.958	457.079
RMSE after normalization	-	0.140	0.138	0.138
Model training time (s)	1.129201	2.341361	2.541292	2.720236
Consumed energy (Joule)	16.938010	35.120412	38.119376	40.803534

Knowing that the consumed energy (Joule)=model training time (seconds) X GPU/CPU Power (Watt). We are using an intel(R) Core (TM) i7-8565U CPU. Its thermal design power is 15 W. Moreover, the operating system is windows 10 64 bits and the Ram is 8192 MB. We notice that contrary to RMSE, the model training time and consumed energy are best for ARIMA and worst for the CNN networks.

4. CONCLUSION

The main contribution of this study is comparing the performance of four prediction models, namely ARIMA a statistical model, MLP a feedforward neural network and LSTM a recurrent neural network and CNN a convolutional neural network. The analysis of the experiment results shows that the convolutional network gives better forecasting results than the LSTM, MLP networks and ARIMA model. However, the model training time and consumed energy are in the opposite order. Meaning, they are best for Arima. This means that to get the better prediction results the CNN or LSTM networks should be used. However, if the model is big, Arima might be a better solution since it has less training time and less consumed energy.

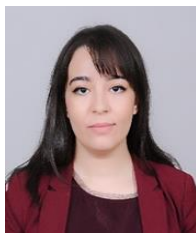
The results of this paper can be further extended in several directions. As a future research direction, it seems interesting to explore the hyper-parameter tuning method knowing that the LSTM is mostly effective in sequence problems and time series allowing information to persist that may lead to better forecasting results. It is also possible to use standard benchmark datasets to validate the proposed prediction models. It is also possible to use other models in the comparative analysis such as the Boltzmann machine and random forests among others.

REFERENCES

- [1] F. Persson and J. Olhager, "Performance simulation of supply chain designs," *Int. J. Prod. Econ.*, vol. 77, no. 3, pp. 231–245, 2002, doi: 10.1016/S0925-5273(00)00088-8.
- [2] P. Hauguel and E. Viardot, "De la supply chain au réseau industriel," *L'Expansion Manag. Rev.*, vol. 101, no. 1, pp. 94–100, 2001.
- [3] B. Fanoodi, B. Malmir, and F. F. Jahantigh, "Reducing demand uncertainty in the platelet supply chain through artificial neural networks and ARIMA models," *Comput. Biol. Med.*, vol. 113, Jan. 2019, Art. no. 103415, doi: 10.1016/j.compbimed.2019.103415.
- [4] L. Ji, Y. Zou, K. He, and B. Zhu, "Carbon futures price forecasting based with ARIMA-CNN-LSTM model," *Procedia Comput. Sci.*, vol. 162, pp. 33–38, 2019, doi: 10.1016/j.procs.2019.11.254.
- [5] M. Matyjaszek, P. R. Fernández, A. Krzemień, K. Wodarski, and G. F. Valverde, "Forecasting coking coal prices by means of ARIMA models and neural networks, considering the transgenic time series theory," *Resour. Policy*, vol. 61, pp. 283–292, Feb. 2019, doi: 10.1016/j.resourpol.2019.02.017.
- [6] M. Ohyver and H. Pudjihastuti, "Arima Model for Forecasting the Price of Medium Quality Rice to Anticipate Price Fluctuations," *Procedia Comput. Sci.*, vol. 135, pp. 707–711, 2018, doi: 10.1016/j.procs.2018.08.215.
- [7] P. J. Oliveira, J. L. Steffen, and P. Cheung, "Parameter Estimation of Seasonal Arima Models for Water Demand Forecasting Using the Harmony Search Algorithm," *Procedia Eng.*, vol. 186, pp. 177–185, 2017, doi: 10.1016/j.proeng.2017.03.225.
- [8] M. S. Al-Musaylh, R. C. Deo, J. F. Adamowski, and Y. Li, "Short-term electricity demand forecasting with MARS, SVR and ARIMA models using aggregated demand data in Queensland, Australia," *Adv. Eng. Informatics*, vol. 35, pp. 1–16, 2018, doi: 10.1016/j.aei.2017.11.002.
- [9] W. Anggraeni, R. A. Vinarti, and Y. D. Kurniawati, "Performance Comparisons between Arima and Arimax Method in Moslem Kids Clothes Demand Forecasting: Case Study," *Procedia Comput. Sci.*, vol. 72, pp. 630–637, 2015, doi: 10.1016/j.procs.2015.12.172.
- [10] M. H. Amini, A. Kargarian, and O. Karabasoglu, "ARIMA-based decoupled time series forecasting of electric vehicle charging demand for stochastic power system operation," *Electr. Power Syst. Res.*, vol. 140, pp. 378–390, 2016, doi: 10.1016/j.epsr.2016.06.003.
- [11] Y. Wang, J. Wang, G. Zhao, and Y. Dong, "Application of residual modification approach in seasonal ARIMA for electricity demand forecasting: A case study of China," *Energy Policy*, vol. 48, pp. 284–294, 2012, doi: 10.1016/j.enpol.2012.05.026.
- [12] R. Sharma and P. Singhal, "Demand forecasting of engine oil for automotive and industrial lubricant manufacturing company using neural network," *Mater. Today Proc.*, vol. 18, pp. 2308–2314, 2019, doi: 10.1016/j.matpr.2019.07.013.
- [13] K. Mason, J. Duggan, and E. Howley, "Forecasting energy demand, wind generation and carbon dioxide emissions in Ireland using evolutionary neural networks," *Energy*, vol. 155, pp. 705–720, 2018, doi: 10.1016/j.energy.2018.04.192.
- [14] I. Matino, S. Dettori, V. Colla, V. Weber, and S. Salame, "Forecasting blast furnace gas production and demand through echo state neural network-based models: Pave the way to off-gas optimized management," *Appl. Energy*, vol. 253, p. 113578, Jun. 2019, doi: 10.1016/j.apenergy.2019.113578.
- [15] E. S. Silva, H. Hassani, S. Heravi, and X. Huang, "Forecasting tourism demand with denoised neural networks," *Ann. Tour. Res.*, vol. 74, pp. 134–154, 2019, doi: 10.1016/j.annals.2018.11.006.

- [16] J. Ke *et al.*, "Hexagon-Based Convolutional Neural Network for Supply-Demand Forecasting of Ride-Sourcing Services," *IEEE Trans. Intell. Transp. Syst.*, vol. 20, no. 11, pp. 4160–4173, 2019, doi: 10.1109/TITS.2018.2882861.
- [17] J. Ke, H. Zheng, H. Yang, and X. (Michael) Chen, "Short-term forecasting of passenger demand under on-demand ride services: A spatio-temporal deep learning approach," *Transp. Res. Part C Emerg. Technol.*, vol. 85, pp. 591–608, Jun. 2017, doi: 10.1016/j.trc.2017.10.016.
- [18] K. Amarasinghe, D. L. Marino, and M. Manic, "Deep neural networks for energy load forecasting," *IEEE Int. Symp. Ind. Electron.*, pp. 1483–1488, 2017, doi: 10.1109/ISIE.2017.8001465.
- [19] I. Slimani, I. El Farissi, and S. Achchab, "Configuration and implementation of a daily artificial neural network-based forecasting system using real supermarket data," *Int. J. Logist. Syst. Manag.*, vol. 28, no. 2, pp. 144–163, 2017, doi: 10.1504/IJLSM.2017.086345.
- [20] A. Said, I. Slimani, and I. El Farissi, "Artificial Neural Networks/or Demand Forecasting: Application Using Moroccan Supermarket Data," *IEEE Int. Conf. Consum. Electron (ICCE)*, Berlin, 2015, pp. 266–271.
- [21] V. A. Profillidis and G. N. Botzoris, "Trend Projection and Time Series Methods," *Modeling of Transport Demand*, pp. 225–270, 2019.
- [22] V. Kotu and B. Deshpande, "Data Science Concepts and Practice," *Morgan Kaufmann*, vol. 2, no. C, 2019.
- [23] T. C. Mills, "ARIMA Models for Nonstationary Time Series," *Appl. Time Ser. Anal.*, pp. 57–69, 2019, doi: 10.1016/b978-0-12-813117-6.00004-1.
- [24] C. C. Aggarwal, "Neural Networks and Deep Learning," *Springer*, 2018.
- [25] S. Hochreiter and J. Urgan Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, pp. 1735–1780, 1997, doi: 10.1162/neco.1997.9.8.1735.
- [26] H. Bousqaoui, S. Achchab, and K. Tikito, "Machine learning applications in supply chains: Long short-term memory for demand forecasting," in *Lecture Notes in Networks and Systems*, vol. 49, pp. 301–317, 2019.
- [27] R. Thanki and S. Borra, "Application of Machine Learning Algorithms for Classification and Security of Diagnostic Images," Elsevier Inc., 2019.
- [28] H. Bousqaoui, I. Slimani, and S. Achchab, "Improving Coordination in Supply Chain Using Artificial Neural Networks and Multi-agent Approach," in *EngOpt 2018 Proceedings of the 6th International Conference on Engineering Optimization*, Springer International Publishing, 2019, pp. 1353–1359.

BIOGRAPHIES OF AUTHORS



Halima Bousqaoui obtained her degree from National Higher School for Computer Science and System Analysis (ENSIAS). And is currently pursuing a PhD degree there. Her work focuses on machine learning algorithms such as feed forward and recurrent neural networks in the supply chain.



Ilham Slimani Ing. & PhD from the National School of Computer Science and System Analysis (ENSIAS). She is currently Professor at the Department of Computer Science at Mohammed First University, Faculty of Science Oujda. Her research and publication interests include machine learning, supply chain management, demand forecasting, road traffic forecasting, intrusion detection of computer networks.



Said Achchab, holder of a PhD in Applied Mathematics from the Mohammedia School of Engineering in Rabat as well as a university habilitation in Business Intelligence from the same institution, he also attended the higher cycle of management at the ENCG of Settat. He is currently Professor of Quantitative Finance, Artificial Intelligence and Risk Management at ENSIAS, Director of the Continuing Education Center of the same school between January 2013 and May 2017 and Head of the Department "Computer Science and Decision Support" since January 2018. He is coordinator of the Specialized Master "Engineering for Sustainable Finance and Risk Management" and of the engineering program "Digital Engineering for Finance". He is the founding President of the African Institute of Fintechs. For more than 15 years, he has been carrying out consulting missions for public and private organizations.