E-ISSN: 2676-6167 | P-ISSN: 2538-5100

Journal of Applied Research on Industrial Engineering

www.journal-aprie.com

J. Appl. Res. Ind. Eng. Vol. 9, No. 3 (2022) 331–353.

Paper Type: Research Paper

A Comparative Study of Demand Forecasting Based on Machine Learning Methods with Time Series Approach

Akbar Abbaspour Ghadim Bonab *

Department of Industrial Engineering, South Tehran Branch, Islamic Azad University, Tehran, Iran; akbarabbaspour@outlook.com.

Citation:

Abbaspour Ghadim Bonab, A. (2022). A comparative study of demand forecasting based on machine learning methods with time series approach. *Journal of applied research on industrial engineering, 9*(3), 331-353.

Received: 03/09/2020 Reviewed: 13/10/2020 Revised: 11/02/2021 Accepted: 02/03/2021

Abstract

Demand forecasting can have a significant impact on reducing and controlling companies' costs, as well as increasing their productivity and competitiveness. But to achieve this, accuracy in demand forecasting is very important. On this point, in the present study, an attempt has been made to analyze the time series related to the demand for a type of women's luxury handbag based on a framework and using machine learning methods. For this purpose, five machine learning models including Adaptive Neuro-Fuzzy Inference System (ANFIS), Multi-Layer Perceptron Neural Network (MLPNN), Radial Basis Function Neural Network (RBFNN), Discrete Wavelet Transform-Neural Networks (DWTNN), and Group Model of Data Handling (GMDH) were used. The comparison of the models was also based on the accuracy of the forecasting according to the values of forecasting errors. The RMSE, MAE error measures as well as the R, correlation coefficient were used to assess the forecasting accuracy of the models. The RBFNN model had the best performance among the studied models with the minimum error values and the highest correlation value between the observed values and the outputs of the model. But in general, by comparing the error values with the data range, it is concluded that the models performed reasonably well.

Keywords: Demand forecasting, Time series, Machine learning.

1 | Introduction

Forecasting can be applied to something that will happen in the future and actually creates a relationship between a firm and its environment. This in turn has a significant impress on planning and decision-making activities in organizations [1]. Ali et al. [2] believe that forecasting is an art and science of discovering the future and in fact predicting the behavior and events of an organization or system in the future. Demand forecasting is used in many subjects, from forecasting demand for electricity, water and energy to demand forecasting for products. By demand forecasting, an organization provides the right plans for future challenges or demands, and plans the necessary steps to deal with them. Demand forecasting does not guarantee the success of a strategy. However, failure to do so will result that the decision to invest, support marketing activities, and allocate other resources be based on unconscious and uncertain assumptions about market requirements, and often such a process will lead to wrong decisions.

Corresponding Author: akbarabbaspour@outlook.com

http://dx.doi.org/10.22105/jarie.2021.246283.1192

So, demand forecasting is extremely helpful for organizations and supply chain managers since it provides a great source of information for planning and decision making [3]. Given this, demand forecasting is one of the essential prerequisites in many aspects of supply chain management such as production planning, inventory control, demand planning, order fulfillment [3], [4]. In this regard, the strength of predictive methods is very important in accurate demand forecasting. As rough or inaccurate estimation of demand is one of the main causes of the bullwhip effect harming the entire supply chain [5]. In the meantime, misconceptions do not stem from a lack of predictive methods, but the performance of forecasting methods can have a significant impact on forecast accuracy. In this respect, classical or statistical demand forecasting models such as linear or polynomial regression techniques and time series models suffer from several weaknesses that undermine their utility for any prediction involving complex and non-linear determinants [6]. Therefore, Machine Learning (ML) models thanks to their training abilities can discover complex models. So, the ML models or their combination with classical methods or meta-heuristic algorithms by acquisition knowledge from experience can improve the forecasting process. Empirical research has demonstrated that ML algorithms for time series prediction provide very competitive results, frequently outperforming classical statistical models [7]. Numerous articles have attempted to provide various frameworks based on ML methods for predicting different types of demand and also provided comparisons between the proposed methods considering based on their performance and accuracy. On this matter, Cankurt and Subasi [8] utilized three ML models including MLR, MLP, and SVR to forecast the demand time series for tourism. The results indicated that the SVR model performed better than other two models. Also, Kandananond [9] employed two machine learning models, Support Vector Machine (SVM) and Artificial Neural Networks (ANN), as well as a classical model, Autoregressive Integrated Moving Average (ARIMA), to predict products demand. The results showed that the SVM model has better performance than the other two models. Hence, in this paper, for demand time series forecasting, ML models will be utilized. *Table 1* shows the studies have done on this subject. The following table groups the prediction approaches in studies into four categories. Based on [10]-[13] the first approach is the use of classical methods, the second approach is the use of single ML methods, the third approach is the use of ensemble models, and the fourth one is the hybrid models. In fact, in ensemble method, a group of base models work independently to finally agree on an output. For example, suppose training several ML models on a training dataset and then their output is a new dataset, which is used as the input of another combiner ML algorithm. The main goal of ensemble method is to generate an optimal predictive model by using of some initial models. On the other side, in hybrid methods, some models work together to forecast one single outcome that no aggregating model exist in it.

By reviewing papers on various topics of demand forecasting, it can be said that most of the materials and methods used to forecast demand in different subjects are similar. So, this issue can be investigated by reviewing various studies in demand forecasting. Here more attention is paid to ML models. Much research has concentrated to increase the efficiency and improve the performance of ML models in prediction. In some studies, researchers have tried to improve prediction results by using ensemble models. In a comparative study, various time series prediction models, including classical, single ML, and ensemble models have been studied. The results showed that the use of ensemble models can increase the performance of models in time series predicting [14]. Qiu et al. [15] presented an ensemble Empirical Mode Decomposition (EMD) based Deep Learning (DL) model for predicting load demand. They compared the results with various other models, which showed the superiority of the proposed model over other models. Also, Kilimci et al. [16] utilized nine different time series models, the SVR model, and DL combined with the boosting ensemble strategy to improve demand forecasting and the results show an improvement in forecasts using the proposed framework.

Table 1. Overview of some papers on demand forecasting.

Reference / Study	Discussed Models	Forecasting Performance Measures						Considered Approaches				Description / Keywords
		MSE	RMSE	MAPE	MASE	MAE	R^2	Classic	Single ML	Ensemble	Hybrid	
[19]	LR, RR, SGDR, AdaBoost, XGBoost LGBM, MLP, DNN-LSTM, DNN-GRU		■			■		■	■	■	■	Demand prediction. (C2C) E-commerce. Advertisements. SOTA.
[3]	HR-ARIMA, ARIMAX ETSX, DLR, SVR, ANN ARIMA, ETS, Theta				■			■	■		■	Demand forecasting. Supply chain. Demand volatility.
[20]	GA-SVR, RBF		■				■		■		■	Demand forecasting. Supply chain.
[5]	MDL, ANN	■									■	Demand forecast. Optimal neural network approach. Surrogate data.
[6]	MLP, ANFIS, SVM, CA, IWO, PSO, GA	■	■			■	■		■		■	Demand prediction based on hybrid artificial intelligence and metaheuristic algorithms. Dairy industry.
[21]	ANFIS	■	■	■	■	■	■		■			Forecasting domestic low cost carrier passenger demand. Air transport.
[22]	GA			■					■			Forecasting domestic low cost carrier passenger demand. Air transport.
[23]	LL, ARIMA, ETS RW ANN, LC			■				■	■			Prediction of customer demands for production planning. Pattern recognition.
[24]	SVM, RNN ARMA, MA, ES, Theta					■		■	■			Traditional and ML-based forecasting.
[25]	ANN, KNN	■				■	■		■			Demand forecasting. Supply chain management.

Table 1. Continued.

Reference / Study	Discussed Models	Forecasting Performance Measures						Considered Approaches				Description / Keywords
		MSE	RMSE	MAPE	MASE	MAE	R^2	Classic	Single ML	Ensemble	Hybrid	
[16]	MA, ARIMA, ES Holt-Winter, Holt-Trend, SVM, SVR, DL(MLFANN)				■			■	■			Improved demand forecasting. Supply chain.
[10]	DT, GBT, RF, LR, SG(LR)		■						■	■		Demand prediction. Machine learning. Stacked generalization.
[26]	ANN-OLMAM, SVM	■		■					■		■	Demand forecasting. Supply chain.
[27]	ANN, ANFIS	■		■					■			Demand forecasting. Supply chain.
[28]	ARMA	Related error term						■				Demand forecasting. Automotive aftermarket. Spare parts.
[29]	RNN, LSSVM, MLR, NN, MA, Trend, Naive	Std. dev				■		■	■			Demand forecasting. Supply chain. Bullwhip effect.
[4]	ARIMA	The accuracy of the developed model was evaluated by comparing the experimental and the simulated sales in the same period						■				Demand forecasting. Time series. Food manufacturing.
[30]	ARMA	Related error term						■				Evolution of ARMA demand in supply chains. Demand propagation.
[9]	SVM, ANN, ARIMA			■				■	■			Time series forecasting. Comparison of time series forecasting models.
[17]	ANN, Bat and Firefly algorithms, Multivariate regression	■		■	MdPAE	■	■		■		■	Air Travel demand. Improving ANN performance using Bat and Firefly algorithms. Comparison of models.

Abbaspour Ghadim Bonab | J. Appl. Res. Ind. Eng. 9(3) (2022) 331-353

Table 1. Continued.

Reference / Study	Discussed Models	Forecasting Performance Measures MSE	RMSE	MAPE	MASE	MAE	R^2	Considered Approaches Classic	Single ML	Ensemble	Hybrid	Description / Keywords
[18]	MLP, ANFIS, LSTM, RRA, GWO, IWO, PSO	■				■	■		■		■	Demand prediction of dairy products. Comparison between various hybrid and pure ML models.
[11]	SVR, CART, LR, SARIMA, ANNs, SVR+LR, SARIMA-MetaFA-LSSVR, SARIMA-PSO-LSSVR	SI	■	■	MaxAE	■	R	■	■	■	■	Comparison of single and ensemble models and hybrid models in energy demand forecasting.
[15]	SVR, ANN, DBN, RF, EDBN, EMD-SVR, EMD-ANN, EMD-RF, ARIMA, TF-ε-SVR-SA, SRSVRCABC		■	■				■	■	■	■	Providing Empirical Mode Decomposition (EMD) based deep learning approach.
[8]	MLR, MLP, SVR			RAE	RRSE				■			Tourism time series demand forecasting.
[31]	SVM, RBF			RME					■			Demand forecasting. Supply chain.
[32]	MA, MARS, ARIMA, GM, SVR, ANN	ME	■	■	MPE	■		■	■			A comparative study of time series forecasting approaches.
Current study	MLPNN, ANFIS, RBFNN, GMDH, DWTNN		■			■	R		■		■	Comparison of machine learning demand forecasting methods with time series approach.

Other models that are used to improve forecasting accuracy can be referred to as hybrid models. For the instance, the combination of ML models with metaheuristic algorithms. Hereof, Mostafaeipour et al. [17] combined ANNs with bat and firefly algorithms and found that using these algorithms increases the rate and speed of neural network adaptation to real data. As well, Goli et al. [18] utilized hybrid ML models to predict dairy products demand. For this end, ANFIS, ANN, and Long Short-Term Memory (LSTM) are used and improved with Runner-Root Algorithm (RRA). Then, a comparison between the combination of these ML models with RRA, Gray Wolf Optimization (GWO), Invasive Weed Optimization (IWO) and Particle Swarm Optimization (PSO) algorithms have been done. The results showed that the RRA optimization algorithm has the highest performance compared to the other three algorithms. As well as, Chou and Tran [11] use single and ensemble ML models, and hybrid models based on time series data to predict energy consumption demand in buildings. They found that ensemble and hybrid models are more accurate than single ML models with 2.7% and 72%, respectively.

The present study attempts to provide a framework for demand time series forecasting and compare the accuracy of employed ML models in forecasting a univariate time series related to a type of women's luxury handbag in a leather products company. Comparison of the models' forecasting accuracy will be based on forecasting errors such as RMSE, MAE and correlation coefficient (R) which represents strengths of relationship between targets (observed values) and outputs (predicted values). The general overview of the procedure of the current study shown in *Fig. 1*. The specific objectives of this study are as follows:

- *Indicating a demand forecasting framework with time series approach.*
- *Demonstrating the use of ML models in demand time series forecasting.*
- *Providing a comparative study of the forecasting accuracy of the ML models.*

The remainder of the paper is organized as follows. Section 2 discusses time series forecasting. Section 3 expresses the material and methods, and investigate employed ML models. In Section 4, the results of the models are presented, and finally Section 5 gives the conclusion of the paper.

2 | Time Series Forecasting

Time series forecasting and analysis has attracted many researchers in recent decades. The main goal in analyzing the time series about a phenomenon is to create a statistical model for time-dependent data based on past information about that phenomenon. This makes it possible to forecast the future of the phenomenon. In other words, time series forecasting is the creation of a model of the past that allows future decisions to be made. Time series forecasting is the act of estimating future events using historical data collected over a fixed period of time and is an important area of forecasting which has gained many attentions from various research areas and in line with its popularity, various models have been introduced for the purpose of producing accurate time series forecasts [33]. Time series forecasting is one of the important components in ML, and has many applications in the real world such as finances, hydrology and healthcare [34]. There are various methods for time series forecasting and generally can be divide into classical and ML methods. Also, in a comprehensive review paper [35], the applications of DL in financial time series forecasting are perused and described and papers in this field are categorize. They also talk about ML methods in time series forecasting. Nowadays, artificial intelligence models are widely used to forecast and retrieve time series. Adeniran et al. [36] believe that predicting with the help of time series is in fact interpreting the sequence and historical path taken by a phenomenon and using the results and knowledge discovered for the near future. Parmezan et al. [7] investigated the most popular algorithms in time series forecasting and as they say there is a limited study on models' performance in complex and highly nonlinear data. Due to the importance of forecasting using time series data, in the current study, this issue has been considered and a framework has been devised for it.

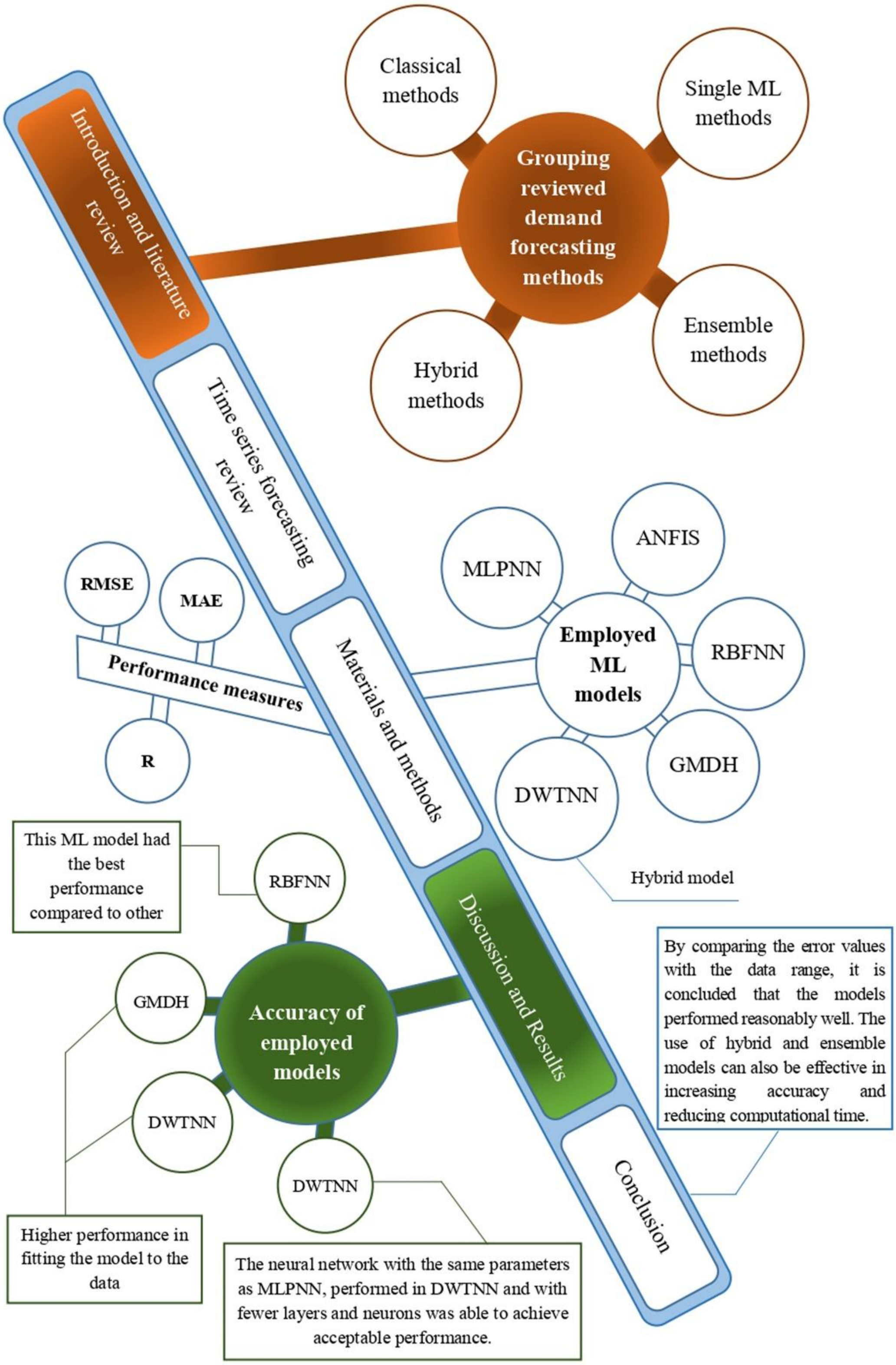

Fig. 1. Demonstration of the procedure utilized on this paper.

3 | Material and Methods

3.1 | Artificial Neural Network (ANN)

The concept of neural networks was introduced [37]. ANNs are a simplified model of the central neural system and like the brain, by processing experimental data transmit the law behind them to the network structure. In fact, the network learns general rules by performing calculations on numerical data, which is why they are called intelligent systems. Neural networks can be used for classification and prediction purposes [38]. Therefore, a neural network can predict the possible future values of a time series based on past values and can be described as a network of simple processing nodes or neurons that are interconnected to each other in a specific order to perform simple numerical manipulations [39]. According to Efendigil et al. [27], ANN has been used in short or long term forecasting of electricity load, energy consumption, tourism and in a wide range of fields with forecast purposes. One of the important areas of application of ANNs is forecasting and recent research activities have shown that ANNs have a high ability to recognize and classify patterns [40]. Modern neural networks are nonlinear statistical tools for data modeling and they are commonly used to model complex relationships between inputs and outputs or to find data patterns [41].

3.1.1 | MLP

One of the most basic ANN models available is the MLP, which mimics the transfer function of the human brain. In this type of neural network, mostly the behavior of human brain networks and signal propagation is considered, and therefore, they are sometimes called Feedforward Networks.

Therefore, in such frameworks, the connection between neurons is Feedforward [42]. MLP usually consist of a set of inputs in one input layer, and one or more hidden layers and an output layer [43]. *Fig. 2* shows the structure of a multilayer perceptron neural network including one input layer, one hidden layers and one output layer. According to Moreno et al. [42] and Lo et al. [44], MLP structure can be explained mathematically by *Eq. (1)*.

$$\hat{y} = f_3\left(\sum_{i=1}^{h} w_{dk} \ . \ f_2\left(\sum_{j=1}^{m} w_{jd} \ . f_1\left(\sum_{i=1}^{n} w_{ij}x_i\right)\right)\right). \tag{1}$$

Where $X = (x_1, \ldots, x_i, \ldots, x_n)$ is the input vector; i, j, d *and* k are indices of samples, input, hidden and output nodes, f_1, f_2 and $f_3$respectively are activation functions of inputs, hidden and output neurons, w_{ij} is the weight of the connection between input i and input neuron j , $w_{jd}$is the weight of the connection between input neuron j and hidden neuron d, w_{dk} is the weight of the connection between hidden neuron d and output neuron k; n, m and h represent the number of samples(inputs), neurons in input layer, neurons in hidden layer, respectively and eventually, $\hat{y}$ is the output provided by output neuron k.

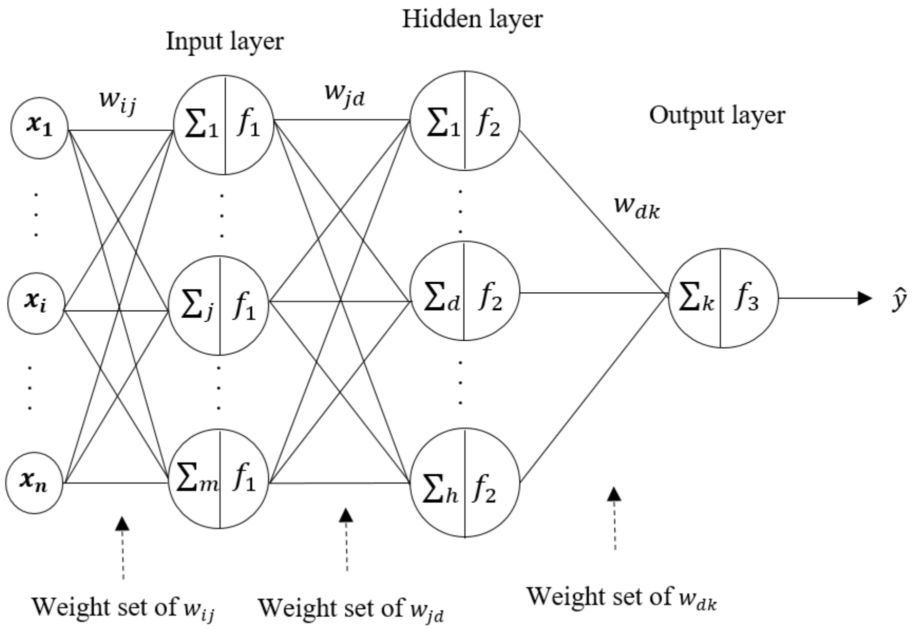

Fig. 2. Multi-layer perceptron neural network including tree layers.

3.1.2 | RBFNN

Similar to the MLP neural network pattern, there are other types of neural networks in which processor units focus on a specific position in terms of processing. This focus is modeled through RBF. In terms of overall structure, RBF neural networks are not much different from MLP networks, and only the type of processing that neurons perform on their inputs is different. Therefore, RBFNs are a special type of feedforward neural networks with radial basis functions used as activation functions [45] and are widely used in regression, classification, pattern recognition, and time series prediction problems. RBFNs often have a faster learning and preparation process. In fact, it will be easier to adjust the neurons because of the focus on specific functional areas. Due to their nonlinear approximation properties, they can model complex relationships [46]. More details and information about neural networks in the following articles [42], [45], [46].

3.2 | GMDH

The GMDH neural network is based on the GMDH algorithm. GMDH is the family of inductive algorithms for multi-parameter data mathematical modeling. The GMDH method was originally formulated to solve for higher-order regression polynomials especially for solving modeling and classification problem [47]. GMDH firstly has presented by Ivakhnenko [48]. One of the most important features of this algorithm is the ability to identify and remove redundant variables [49], [50]. To find the best solution, the GMDH model has several solutions, called partial models or partial derivatives. The coefficients of this model are estimated by the least squares method. This algorithm gradually increases the number of partial models and selects a structural model with the desired complexity. This process is called the self-organizing model. Therefore, it can be said GMDH is a self-organized model consisting of several layers, and each layer is made up of several neurons [51]. The most popular basic function used in GMDH is the Kolmogorov-Gabor polynomial defined by *Eq. (2)*.

$$y = a_0 + \sum_{i=1}^{m}(a_i x_i) + \sum_{i=1}^{m}\sum_{j=1}^{m}\left(a_{ij}x_i x_j\right) + \sum_{i=1}^{m}\sum_{j=1}^{m}\sum_{k=1}^{m}\left(a_{ijk}x_i x_j x_k\right) + \cdots. \tag{2}$$

In which $X = (x_i, x_j, x_k \dots, x_m)$ Indicates the vector of input variables, m represent the number of input, y is output variable, and $A = (a_0, a_i, a_{ij,} a_{ijk}, \dots, a_m)$ is the vector of coefficients of the model to be optimized [47] and [52]. The Kolmogorov-Gabor polynomial can approximate any stationary random sequence of observations and can be computed by either adaptive methods or a system of Gaussian normal equation [50].

3.3 | ANFIS

The combination of Fuzzy Inference Systems (FIS) which are based on logical rules and ANN method that has the ability to extract knowledge from numerical information, led to the introduction of ANFIS by Jang [53]. These networks, while having the ability to learn neural networks and the inference power of fuzzy systems, have the power to find any type of nonlinear model and can accurately link inputs (initial values) to outputs (predicted values). To show the ANFIS structure which is based on Takagi–Sugeno fuzzy inference system [54], for a system with two input (x, y), a logical AND operation and one output $(\hat{y})$ we use two fuzzy IF-THEN control rules specified by *Eqs. (3)* and *(4)*.

$$\text{IF x is } A_1 \text{ AND y is } B_1 \text{ THEN } f_1 = p_1 x + q_1 y + r_1. \tag{3}$$

$$\text{IF x is } A_2 \text{ AND y is } B_2 \text{ THEN } f_2 = p_2 x + q_1 y + r_2. \tag{4}$$

Where A_i and B_i $(i = 1, 2)$ are fuzzy sets of input variables x and y, and p_i, q_i, r_i $(i = 1, 2)$ design parameters identified during training process [55].

After entering x *and* y into the fuzzy sets, a series of membership degrees come out as O_i^1 or outputs of layer 1 as *Eq. (5)*.

$$O_i^1 = \mu_{A_i}(x) \,.\quad i = 1,2 \qquad O_i^1 = \mu_{B_i}(y) \,.\quad i = 1,2 \tag{5}$$

The fuzzy membership function of the crisp input variables x *and* y is determined in layer 1. This layer also is called the membership layer or fuzzification layer. Where μ_{A_i} and μ_{B_i} are membership functions for fuzzy sets A_i and B_i.

The nodes of the second layer also called rule nodes, act as fixed operators and multiply inputs (outputs of the first layer) and express the intensity of activation of the corresponding rules as outputs, represented at *Eq. (6)* [56]:

$$O_i^2 = w_i = \mu_{A_i}(x) \times \mu_{B_i}(y) \,.\quad i = 1,2 \tag{6}$$

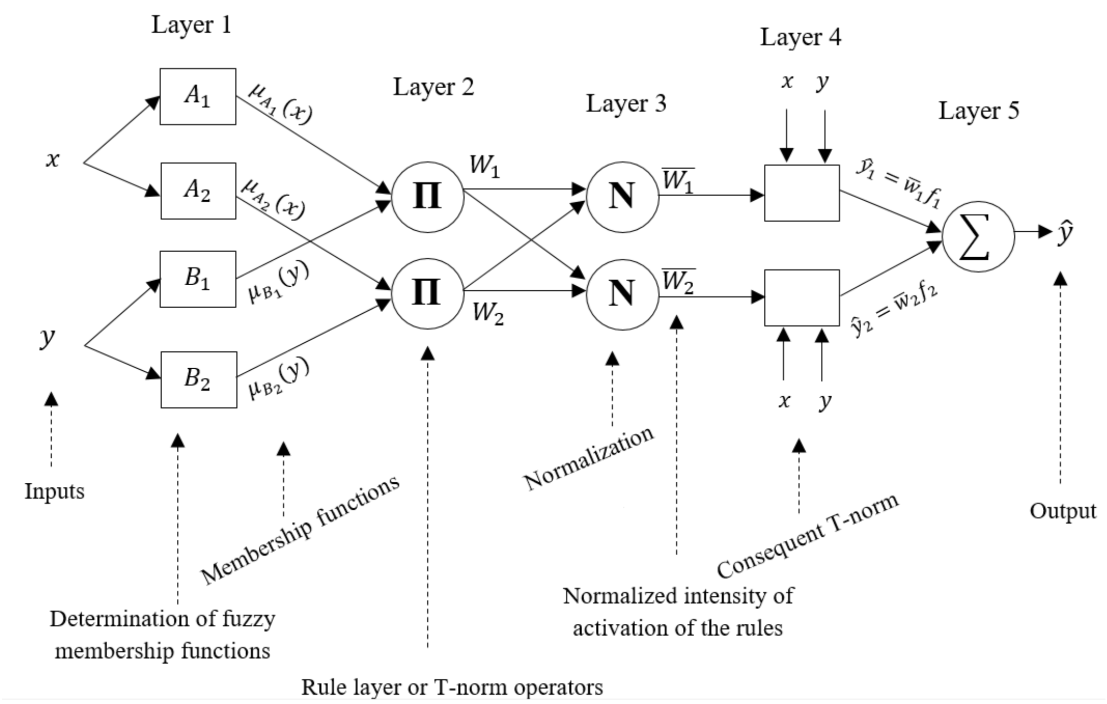

Fig. 3. ANFIS structure with two input variable and two rules.

Layer 3 nodes normalize the intensity of activation of the rules defined by *Eq. (7)*. In other hand, in fact, each node in this layer calculates the correctness ratio of each fuzzy rule to the correctness sum of all the rules and returns it as the output.

$$O_i^3 = \overline{W_i} = \frac{W_i}{W_1 + W_2} \,.\qquad i = 1,2 \tag{7}$$

Layer 4. Each node in this layer has a node function as *Eq. (8)*. Where W_i is the output of previous layer and, p_i, q_i, r_i represent the set of parameters. The parameters of this layer are usually called conclusion parameters because the network output is linear concerning these parameters.

$$O_i^4 = \hat{y}_i = \overline{W}_i(p_i x + q_i y + r_i). \quad i = 1,2 \tag{8}$$

Finally, layer 5. output node, the only node in this layer and its output is the sum of all its inputs. In fact, it doses summation of all received inputs, defined by *Eq. (9)*.

$$O_i^5 = \hat{y} = \sum_i \overline{W}_i f_i = \frac{\sum_i \overline{w}_i f_i}{\sum_i \overline{w}_i} \,.\qquad i = 1,2 \tag{9}$$

ANFIS trains and updates all parameters including MF parameters and p_i, q_i, r_i by using the gradient descent method and the least square methods.

Abbaspour Ghadim Bonab | J. Appl. Res. Ind. Eng. 9(3) (2022) 331-353

3.4 | DWTNN

Wavelet Networks (WNs) are a new class of networks which have been used with great success in a wide range of applications [57]. In literature, integrating ANN with wavelet transform in order to reinforce its modelling performance [58]. Kærgaard et al. [59] used DWTNN hybrid model and compared it with several other hybrid models, which proved the superiority of this model. In the model used in this paper, first a wavelet analysis is applied to the data and then this data is used as the input of the neural network. In fact, wavelet transform can better extract those properties in the data, which is not possible for the neural network. The general structure of the model used is shown in *Fig. 4* with two levels of decomposition. In this figure, a_i indicates low frequency, better for long term prediction and d_i indicates high frequency, better for short term prediction. However, the use of these signals here is the responsibility of the neural network. Discussion about the details of this model is out of the scope of this article, so refer to papers [57] and [58] for more information.

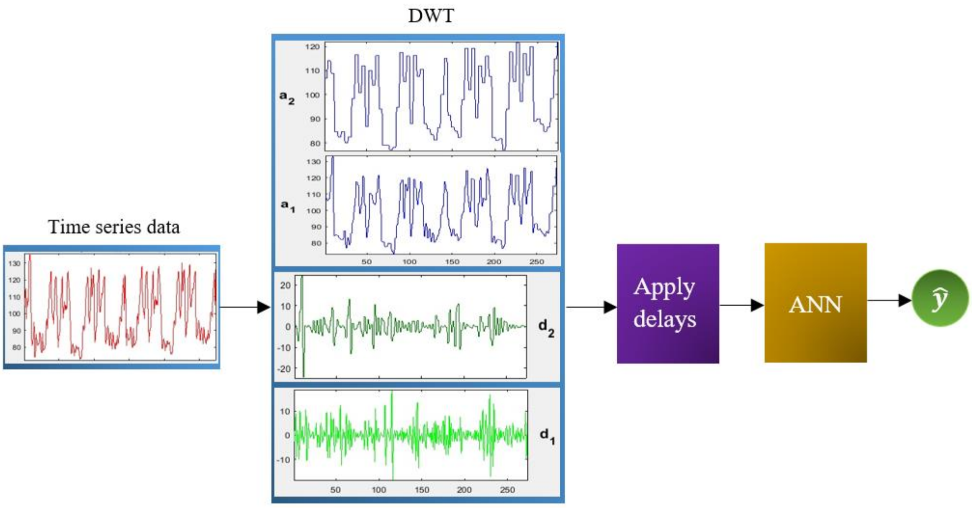

Fig. 4. General structure of the employed DWTNN model.

3.5 | Structure of the Proposed Model

When we make predictions, the assumption is implicit that behavior is almost predictable over time, and that there is not much change in the behavior of the system. This is why we can do predictive action, because if a phenomenon is constantly changing, it will not be predictable or the prediction will be very hard. Therefore, it can be claimed that based on the behavior of the system in the past and the study of information taken from the past, the future of the system can be forecasted. So the important thing is having enough data from the past of the system and in the same past, creating a relationship between the past and the future through the same data. Then, since this relationship describes the past well and the system has not changed during this time, we can claim that the current state of the system and the future can be predicted. To do this, we need to make sure that ML can incorporate the dynamics of the data into the model, because under normal circumstances it may not be able to extract these dynamics from the data. So, we need to consider the time periods that have a jump in demand and bring this trend to the modeling network. For example, if we consider this week as x_t then x_{t-5} means 5 weeks ago or x_{t-7} means 7 weeks ago which can play a role in today's demand. In fact, by doing so, we bring our knowledge to the network. Of course, input (feature) selection methods can also be used for this purpose. In fact, to model a time series, a relationship must be established between the current value of the time series and the value of the same series in the past. In this regard, the neural network or any other modeling system is used to estimate this relationship. In fact, this system is not without inputs, but inputs are made up of past (delayed) values. Time information can be included through a set of delays that are added to the input, so that by using this method we can demonstrate data at different points in time. So the smart choice of delays will have a big impact on the accuracy and functionality of the predictive model. We must first determine the delays so that estimates can be made based on them. For example, here we consider the delays as 1, 2, 3, 4, 6, 8, 12, 20. This means that if these numbers represent weeks, then knowing what the value of our time series was

1, 2, 3, 4, 6, 8, 12 and 20 weeks ago, then we can predict the amount of time series on the forecast horizon, which here is equal to the smallest delay which here is one week. It should be noted that the size of the data is affected by the largest delay and the size of the forecast horizon is affected by the smallest delay. Actually, as much as maximum delay, we lose the initial values of the data (*Fig. 5*). So the size of the delays should be proportional to the size of the data. *Fig. 6* depicts the process of demand forecasting using time series data.

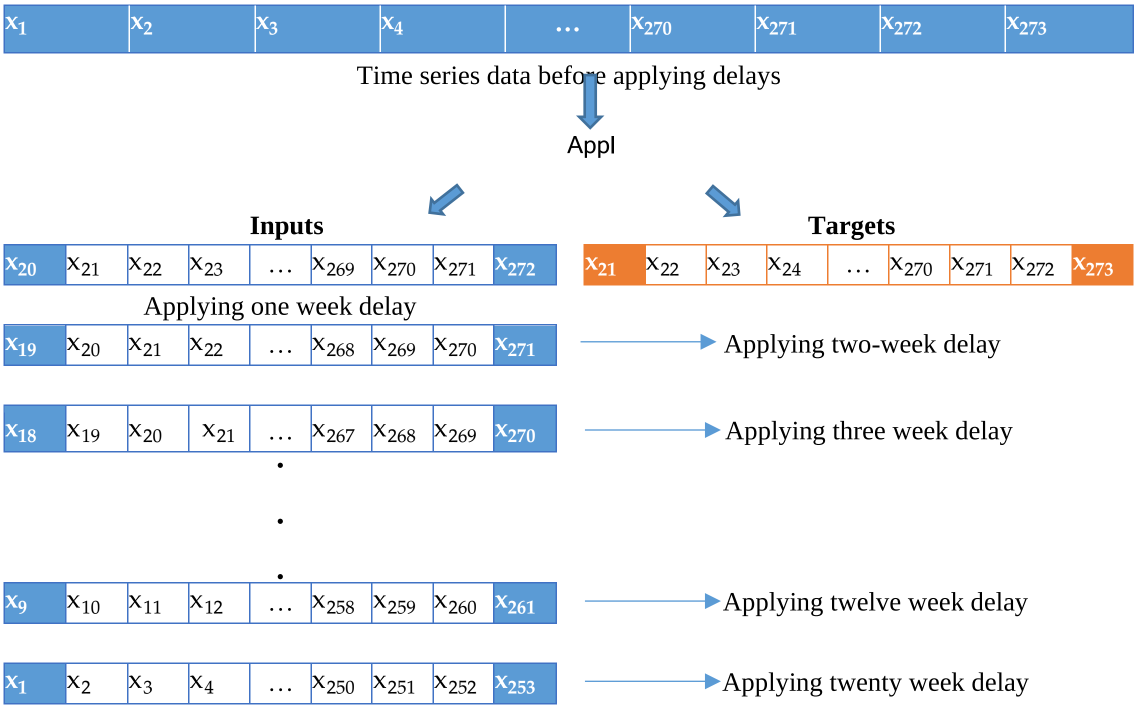

Fig. 5. Applying delays on time series data.

3.6 | Performance Evaluation Criteria

A number of criteria are used to evaluate the performance and accuracy of forecasting models. For this purpose, Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) were employed. Also, correlation coefficient used to represent strength between outputs and targets. RMSE is the standard deviation of the difference between outputs and targets (forecasted and observed values). MAE demonstrates the average of the absolute difference between the outputs and targets value. Where n is the number of samples, if x_i represents the targets (observed values) and $\hat{y}_i$ represents the outputs (predicted values), then RMSE and MAE can be expressed as *Eqs. (10)* and *(11)* [60], [61].

$$\text{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i - \hat{y}_i)^2}. \tag{10}$$

$$\text{MAE} = \frac{1}{n}\sum_{i=1}^{n}|x_i - \hat{y}_i|. \tag{11}$$

Both RMSE and MAE express the average magnitude of the prediction error in the unit of dependent Variable and their range can vary from 0 to ∞. However, the best value for RMSE and MAE depends on the domain of the data and the purpose of the problem, but in general, the lower the value, the better. As these measures are scale dependent, so they can be used to compare different methods on the same data set [62].

Abbaspour Ghadim Bonab | J. Appl. Res. Ind. Eng. 9(3) (2022) 331-353

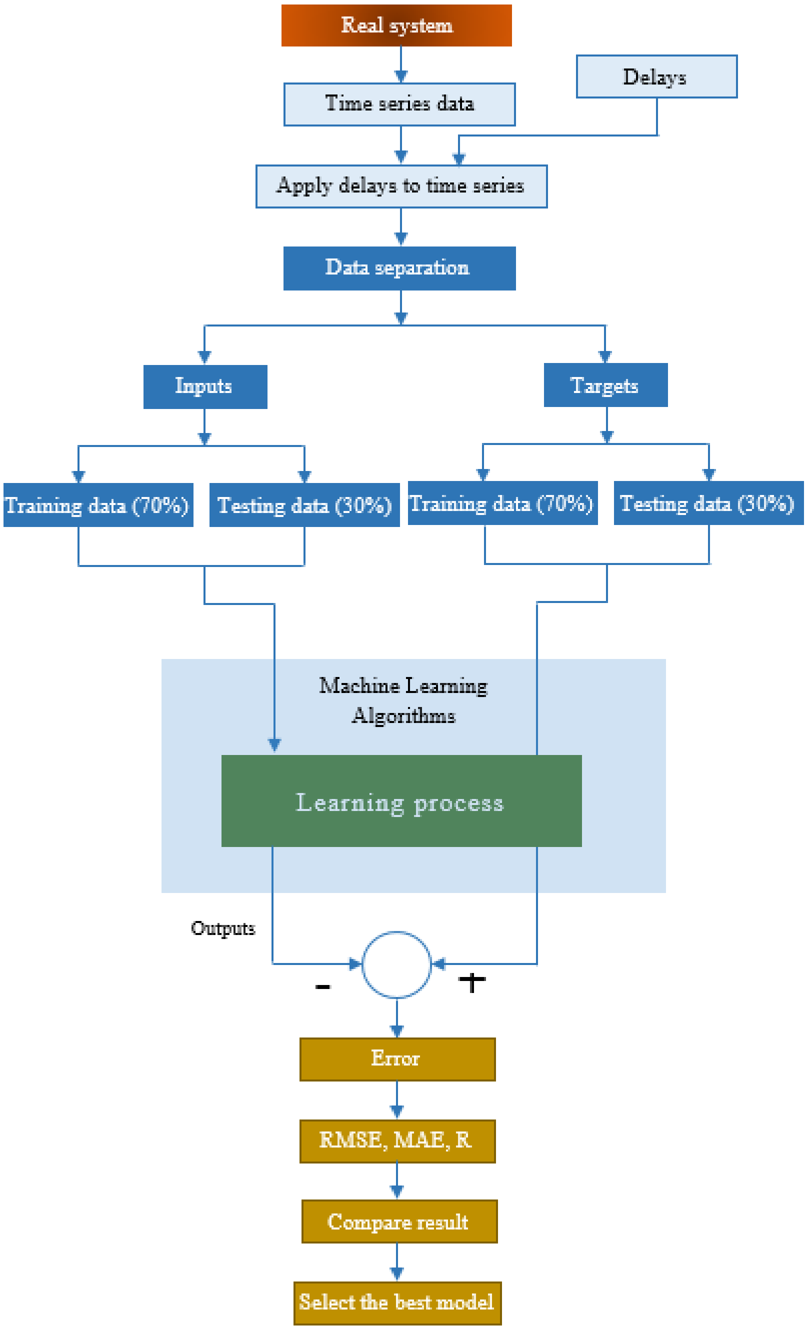

Fig. 6. The framework of demand forecasting using time series data.

Eq. (12) represents the degree of relationship between outputs and targets. The value of R is between -1 and +1, when *R=0* indicates no correlation or linear relationship between outputs and targets (there isn't any link between values), *R = 1* demonstrates a total positive linear correlation or linear relationship between outputs and targets and *R=-1* represents a perfect negative correlation [63]. Where $\bar{x}$ and $\bar{y}$ represent mean of observed values (targets) and mean of predicted values (outputs) respectively.

$$R = \frac{\sum_{i=1}^{n}(x_i - \bar{x})(\hat{y}_i - \bar{y})}{\sqrt{\sum_{i=1}^{n}(x_i - \bar{x})^2 \sum_{i=1}^{n}(y_i - \bar{y})^2}}. \quad (12)$$

3.7 | Case Study

The analyzed dataset is related to the time series of demand for a type of women's luxury handbag in a Luxury leather products company. The data are related to the demand for the last six years, which have been collected as a weekly demand. It should be noted that demand fluctuations in data are related to discount periods or increased demand over specific time periods.

4 | Results and Discussion

The process of demand forecasting will utilize developed ML models based on collected time series data. The forecast horizon, as mentioned in Section 3.5, will be short-term, equal to 1 week. Of course, it is worth noting that based on the richness of the data, it will be possible to increase the forecast horizon. In all models, 70% of the data will be used as training data and the rest as test data and, if needed, validation. It should be noted that in the result table and figures, all data includes training and test sets. Also, all models have been developed in MATLAB software.

4.1 | ANFIS

In ANFIS modeling, three methods including grid partitioning, subtractive clustering and FCM, can be used to generate an initial FIS. In this paper, FCM has been used for this purpose. FCM is a clustering-based method and uses fuzzy c-means. The parameters involved in generating the initial FIS are considered as follows: number of clusters =10, partition matrix exponent = 2 and maximum number of iterations= 200. In the ANFIS model, there are a number of parameters in the training phase to minimize the amount of error function. In fact, by setting these parameters, we change the model in such a way that the error has the least value. In this regard, parameters including minimum improvement=1e-5, maximum number of epochs = 300, error goal = 0, initial step size = 0.01, step size decrease rate =0.9, step size increase rate = 1.1 and hybrid optimization method are considered as mentioned. The performance of the proposed ANFIS model was determined using RMSE and MAE (*Table 2*). Also, *Fig. 7* indicates the comparison between targets and outputs of the ANFIS model. And according to that, the model has been able to have relatively good results and has done a good forecast, except for a few cases where demand has suddenly increased or decreased. It should be noted that all data includes train and test data. *Fig. 8* demonstrates correlations between targets and outputs.

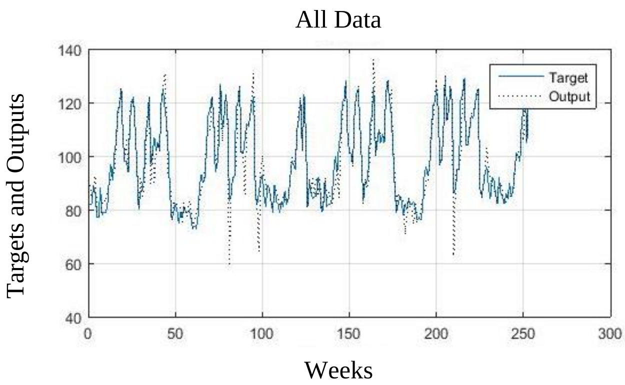

Fig. 7. Comparison of targets and outputs for ANFIS model.

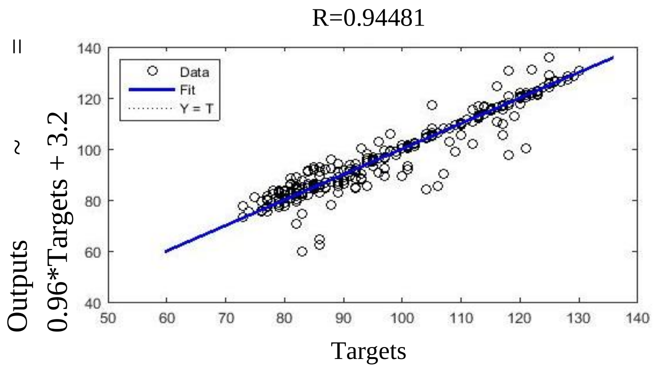

Fig. 8. Correlation between the targets and outputs for ANFIS model.

Abbaspour Ghadim Bonab | J. Appl. Res. Ind. Eng. 9(3) (2022) 331-353

4.2 | GMDH

To make forecasting with the GMDH model, after determining the training and testing data, the optimal values of the parameters involved in the implementation of the GMDH model were set during several different stages to get the best results. Finally, the parameters including the maximum number of neurons=80, the maximum number of layers = 6, and selection pressure = 0.6 were used with the values mentioned. According to the above parameters, GMDH performs calculations using 80 neurons and finally stops in the sixth layer. Also, to prevent overgrowth and complexity, the GMDH algorithm only selects values in each layer that have good predictive power and uses the s_p (selection pressure) parameter for this purpose. s_p=0.6 means that the GMDH model removes 60% of values and selects the top 40% in each layer. The accuracy of the forecast and the performance errors including RMSE and MAE can be found in the *Table 2*. *Fig. 9* indicates comparison of targets and outputs for GMDH model. Also, *Fig. 10* demonstrates the correlation between targets and outputs, and accordingly, it can be said that the model performed better in predicting lower demand.

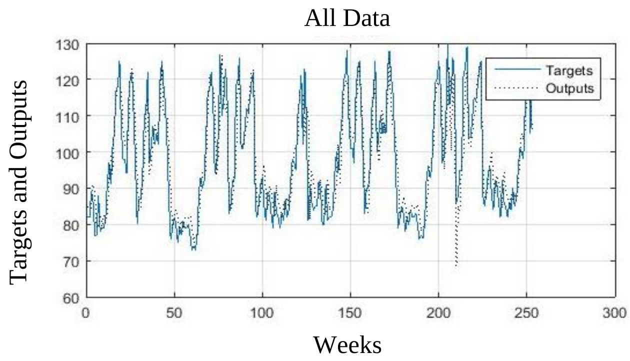

Fig. 9. Comparison of targets and outputs for GMDH model.

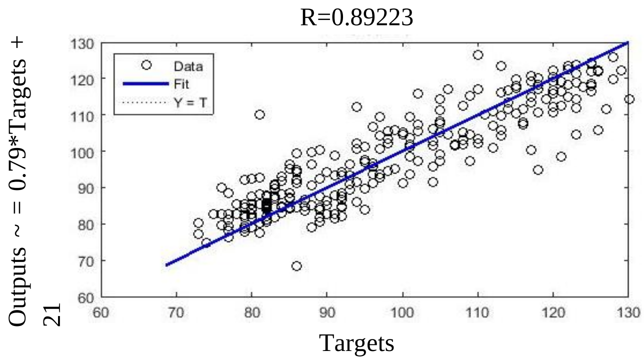

Fig. 10. Correlation between the targets and outputs for GMDH model.

4.3 | MLPNN

In this section, a three-layer neural network was used to predict demand, including an input layer with twenty neurons, a hidden layer with ten neurons, and an output layer including one neuron. It is worth mentioning that MATLAB considers all layers except the output layer as hidden layers. The developed model was trained using the Levenberg-Marquardt backpropagation algorithm, known as trainlm function in MATLAB. Also, the transfer functions used in the hidden layers were tansig and in the output layer, it was Purelin. The performance of the model in terms of RMSE and MAE can be seen in *Table 2*, and according to it, it can be claimed that the model has acceptable accuracy and performance. *Fig. 11* shows the performance of the model in forecasting demand (outputs) compared to targets. As well as, *Fig. 11* indicates that the model performed well in forecasting strong demand fluctuations and it can recognize the pattern created by the targets (observed data) in an acceptable way. *Fig. 12* demonstrates the correlation between targets and outputs.

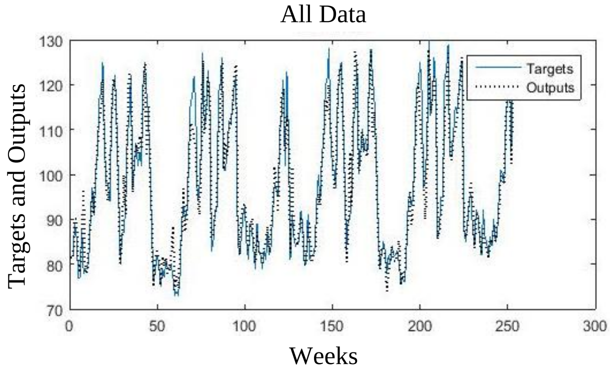

Fig. 11. Comparison of targets and outputs for MLPNN model.

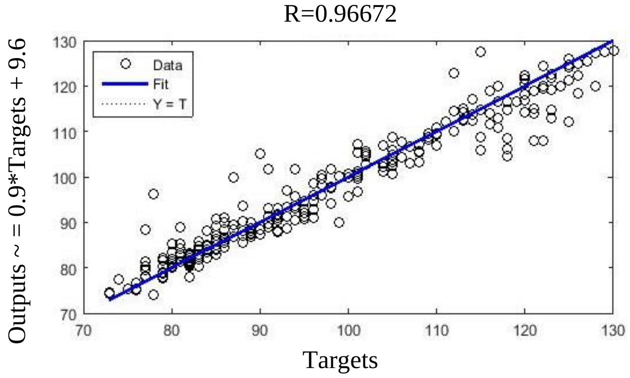

Fig. 12. Correlation between the targets and outputs for MLPNN model.

4.4 | RBFNN

There are parameters involved in the design of an RBF neural network that must be determined before network implementation. Goal, what we expect from the network to reach this amount or is the expected MSE error which is zero by default in MATLAB. Spread, which represents the scatter of RBF functions, and by default in MATLAB is equal to one. The next parameter is maximum number of neurons (MN), which by default is equal to the number of inputs. That is, the RBF neural network up to the MN increases the number of neurons, and if the error reaches the goal value, it will stop the process before reaching the MN. The RBF model was implemented with different parameter values over several stages. The best results were obtained with Goal = 0, Spread = 40, MN = 177. *Fig. 13* indicates the comparison between Targets and Outputs of the RBFNN model for all data and According to that, it can be said that the model has been able to make a good Approximation. Performance accuracy for RBFNN model can be found in *Table 2*. *Fig. 14* represents relationship between targets and outputs with R=0.99176 and according to that, the model was able to cover the prediction of low and high demands well.

Abbaspour Ghadim Bonab | J. Appl. Res. Ind. Eng. 9(3) (2022) 331-353

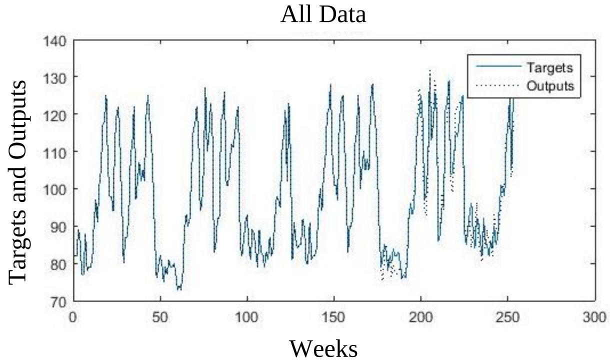

Fig. 13. Comparison of targets and outputs for RBFNN model.

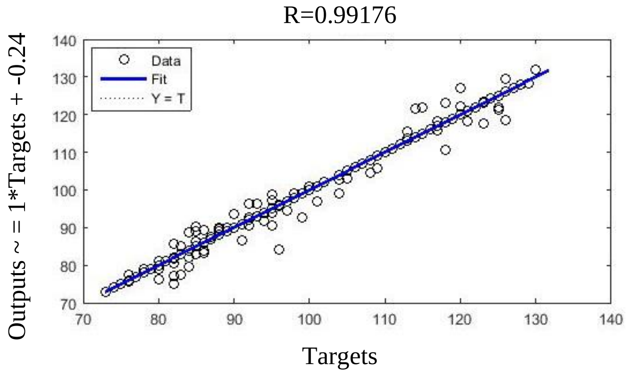

Fig. 14. Correlation between the targets and outputs for RBFNN model.

4.5 | DWTNN

In the combination of the wavelet transform and the neural network, input data for the neural network is actually provided by the wavelet transform. That is, the data is passed through a series of filters, and in fact gives the neural network more details and more complete information. The wavelet decomposes and separates features that the neural network cannot do on its own. The basis of this work is by dividing the signal (time series data) into high-frequency components and low-frequency components. In fact, with this analysis, the privilege is provided for the neural network to select data of high importance. Normally, the number of network inputs was 8, the same size as the number of delays, but with the use of wavelet transform, the number of inputs increased to 32. In fact, by doing this, we increase the information and make the work easier for the neural network. Based on this, the neural network with the same parameters as MLPNN, performed in DWTNN and with only two layers and 3 neurons was able to achieve acceptable performance. Performance of the model can be found in *Table 2*. *Fig. 15* indicates the comparison of the targets and outputs. Also *Fig. 16* demonstrates the correlation between targets and outputs with R=.94727.

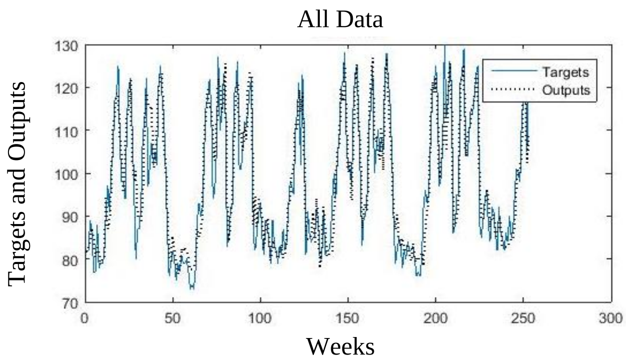

Fig. 15. Comparison of targets and outputs for DWTNN model.

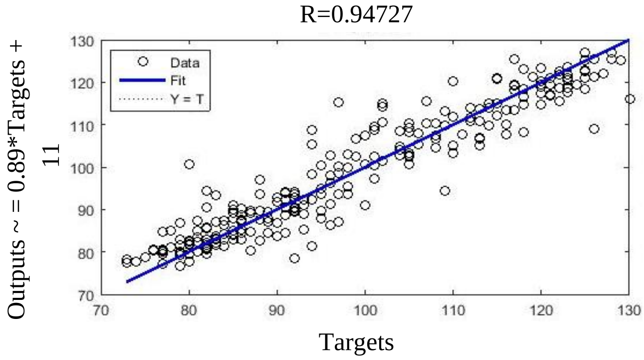

Fig. 16. Correlation between the targets and outputs for DWTNN model.

4.6 | Performances of the Employed Models

Five models were used in demand forecasting. The models used were ML models that predicted the time series of a product's demand. The forecast interval in all models was one week, which is a short period for this time series. Almost all models had acceptable performance in forecasting considering sharp fluctuations in demand at different intervals due to the discounts on products or specific intervals when product demand is higher. The performance of the models based on RMSE, MAE and R is given in *Table 2* and *Table 3*. Accordingly, the RBFNN model has the lowest values of RMSE and MAE for training and test sets, and maximum value for R. So, RBFNN performed better than other models. Also, the closeness of the RMSE values in the training and test sets indicates the good performance in fitting the model to the data and overcoming the noise in the data. Of this, considering the *Fig. 17*, GMDH and DWTNN have better performance. Considering that RMSE and MAE are scale-dependent measures, then it can be said in the scale of our data, models have acceptable performance. Also, *Fig. 18* and *Fig 19* depict the MAE and R for employed models, respectively.

Table 2. The performance of the employed models.

Model	Training		Test		All Data	
	RMSE	MAE	RMSE	MAE	RMSE	MAE
ANFIS	3.1001	1.7766	8.3233	6.0639	5.2473	3.0645
GMDH	6.6743	5.1841	7.8107	6.1053	7.035	5.4608
MLPNN	2.6363	1.8667	6.4037	4.9501	4.0415	2.6586
RBFNN	6.777e-12	4.8973e-12	3.6829	2.9066	2.0186	0.87314
DWTNN	4.2695	3.0399	6.0985	4.5022	4.998	3.6163

Table 3. Strength of the relationship between targets and outputs.

Model	Training	Test	All Data
	R	R	R
ANFIS	0.98048	0.86343	0.94481
GMDH	0.90553	0.85607	0.89223
MLPNN	0.98726	0.91009	0.96672
RBFNN	1	0.97481	0.99176
DWTNN	0.96325	0.9268	0.94727

Abbaspour Ghadim Bonab | J. Appl. Res. Ind. Eng. 9(3) (2022) 331-353

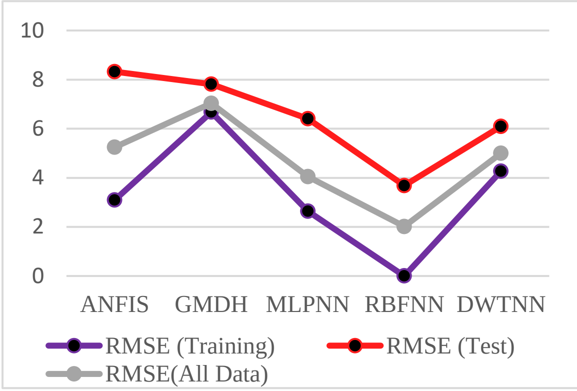

Fig. 17. RMSE values for training and test sets.

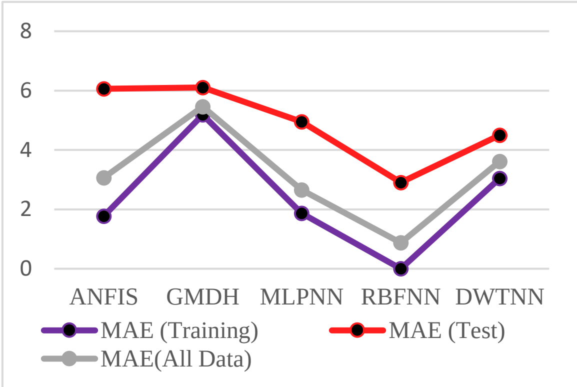

Fig. 18. MAE values for training and test set.

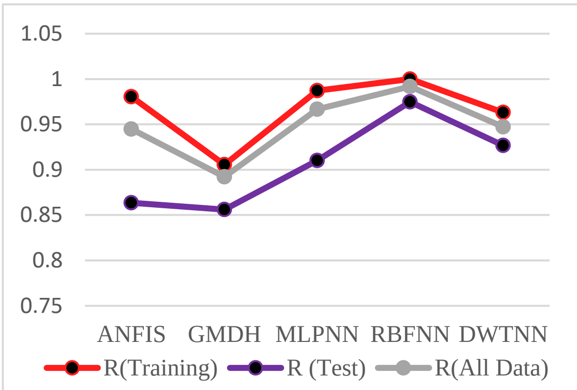

Fig. 19. Correlation between targets and outputs.

5 | Conclusion

Demand and subsequently sales are a vital artery of a business. What helps a business pay employees, cover operating costs, increase inventory, offer new products, and attract new investment. Demand forecasting is an important part of a business's financial planning and is a self-assessment tool that uses past and present demand statistics to intelligently predict future performance. With a careful forecast of demand, businesses can plan for their future. So accurate forecasting is one of the biggest challenges facing companies. Because predicting more than required demand or less than the required demand will cause a lot of costs for the businesses, such as inventory costs or shortage costs. Therefore, it is essential to use the right framework as well as to use models that can accurately predict demand. Hereof, in the present study, using a framework, the time series related to the demand for a women's luxury handbag in a leather products company was analyzed and predicted with five different ML models. Of all the models, the RBFNN had the minimum forecasting errors including RMSE and MAE, and also the highest correlation between targets and outputs. So, the RBFNN model performed best among other models but, the other models also performed well with a slight difference and had acceptable performance. As well as, the DWTNN hybrid model was also able to provide acceptable results with fewer neurons and layers compared to pure MLP. Hence, the use of ML models that have the ability to learn and predict complex problems, as a tool in hand can play a significant role in the growth and prosperity of companies.

Suggestions for future research:

- *As future research, other techniques, algorithms, and methods can be employed to more improve the results.*
- *In the field of ensemble and hybrid models, there is a lot of maneuverability, so this opportunity can be used to provide models with higher accuracy.*

Management insights:

- *It is expected that by using demand forecasting models, the management teams of companies and organizations can have an accurate estimate of future demand and subsequently be able to make the right decisions.*
- *Also, by recognizing the behavior of the system in the future, organizations can identify paths of development and progress, so this will cause their growth and prosperity.*

References

[1] Darvishinia, R., Ebrahimzadeh Shermeh, H., & Barzkar, S. (2019). Development of a forecasting model for investment in Tehran stock exchange based on seasonal coefficient. *Journal of applied research on industrial engineering*, *6*(4), 333-366. https://dx.doi.org/10.22105/jarie.2019.1963 92.1103

[2] Ali, A. Y., Hassen, J. M., & Wendim, G. G. (2019). Forecasting as a framework for reducing food waste in Ethiopian university canteens. *Journal of applied research on industrial engineering*, *6*(4), 374-380. https://dx.doi.org/10.22105/jarie.2020.206803.1109

[3] Abolghasemi, M., Beh, E., Tarr, G., & Gerlach, R. (2020). Demand forecasting in supply chain: the impact of demand volatility in the presence of promotion. *Computers & industrial engineering*, *142*, 106380. https://doi.org/10.1016/j.cie.2020.106380

[4] Fattah, J., Ezzine, L., Aman, Z., El Moussami, H., & Lachhab, A. (2018). Forecasting of demand using ARIMA model. *International journal of engineering business management*, *10*, 1847979018808673. https://doi.org/10.1177/1847979018808673

[5] Lau, H. C., Ho, G. T., & Zhao, Y. (2013). A demand forecast model using a combination of surrogate data analysis and optimal neural network approach. *Decision support systems*, *54*(3), 1404-1416. https://doi.org/10.1016/j.dss.2012.12.008

[6] Goli, A., Zare, H. K., Moghaddam, R., & Sadeghieh, A. (2018). A comprehensive model of demand prediction based on hybrid artificial intelligence and metaheuristic algorithms: a case study in dairy industry. *Journal of industrial and systems engineering*, *11*(4), 190-203. http://www.jise.ir/article_76524.html

[7] Parmezan, A. R. S., Souza, V. M., & Batista, G. E. (2019). Evaluation of statistical and machine learning models for time series prediction: identifying the state-of-the-art and the best conditions for the use of each model. *Information sciences*, *484*, 302-337. https://doi.org/10.1016/j.ins.2019.01.076

[8] Cankurt, S., & SUBAŞI, A. (2016). Tourism demand modelling and forecasting using data mining techniques in multivariate time series: a case study in Turkey. *Turkish journal of electrical engineering and computer sciences*, *24*(5), 3388-3404. https://doi.org/10.3906/elk-1311-134

[9] Kandananond, K. (2012). A comparison of various forecasting methods for autocorrelated time series. *International journal of engineering business management*, *4*, 4. https://journals.sagepub.com/doi/pdf/10.5772/51088

[10] Tugay, R., & Öğüdücü, Ş. G. (2017). Demand prediction using machine learning methods and stacked generalization. *Proceedings of the 6th international conference on data science, technology and applications*, 1, 216–222. https://doi.org//10.5220/0006431602160222

[11] Chou, J. S., & Tran, D. S. (2018). Forecasting energy consumption time series using machine learning techniques based on usage patterns of residential householders. *Energy*, *165*, 709-726. https://doi.org/10.1016/j.energy.2018.09.144

[12] Ardabili, S., Mosavi, A., & Várkonyi-Kóczy, A. R. (2019, September). Advances in machine learning modeling reviewing hybrid and ensemble methods. *International conference on global research and education* (pp. 215-227). Springer, Cham.

[13] Shih, H., & Rajendran, S. (2019). Comparison of time series methods and machine learning algorithms for forecasting Taiwan Blood Services Foundation's blood supply. *Journal of healthcare engineering*, *2019*. https://doi.org/10.1155/2019/6123745

[14] Pavlyshenko, B. M. (2019). Machine-learning models for sales time series forecasting. *Data*, *4*(1), 15. https://doi.org/10.3390/data4010015

[15] Qiu, X., Ren, Y., Suganthan, P. N., & Amaratunga, G. A. (2017). Empirical mode decomposition based ensemble deep learning for load demand time series forecasting. *Applied soft computing*, *54*, 246-255. https://doi.org/10.1016/j.asoc.2017.01.015

[16] Kilimci, Z. H., Akyuz, A. O., Uysal, M., Akyokus, S., Uysal, M. O., Atak Bulbul, B., & Ekmis, M. A. (2019). An improved demand forecasting model using deep learning approach and proposed decision integration strategy for supply chain. *Complexity*, *2019*. https://doi.org/10.1155/2019/9067367

[17] Mostafaeipour, A., Goli, A., & Qolipour, M. (2018). Prediction of air travel demand using a hybrid artificial neural network (ANN) with Bat and Firefly algorithms: a case study. *The journal of supercomputing*, *74*(10), 5461-5484. https://doi.org/10.1007/s11227-018-2452-0

[18] Goli, A., Moeini, E., Shafiee, A. M., Zamani, M., & Touti, E. (2020). Application of improved artificial intelligence with runner-root meta-heuristic algorithm for dairy products industry: a case study. *International journal on artificial intelligence tools*, *29*(05), 2050008. https://doi.org/10.1142/S0218213020500086

[19] Rai, S., Gupta, A., Anand, A., Trivedi, A., & Bhadauria, S. (2019, July). Demand prediction for e-commerce advertisements: a comparative study using state-of-the-art machine learning methods. *2019 10th international conference on computing, communication and networking technologies (ICCCNT)* (pp. 1-6). IEEE. DOI: 10.1109/ICCCNT45670.2019.8944783

[20] Guanghui, W. A. N. G. (2012). Demand forecasting of supply chain based on support vector regression method. *Procedia engineering*, *29*, 280-284. https://doi.org/10.1016/j.proeng.2011.12.707

[21] Srisaeng, P., Baxter, G. S., & Wild, G. (2015). An adaptive neuro-fuzzy inference system for forecasting Australia's domestic low cost carrier passenger demand. *Aviation*, *19*(3), 150-163. https://doi.org/10.3846/16487788.2015.1104806

[22] Srisaeng, P., Richardson, S., Baxter, G., & Wild, G. (2016). Forecasting Australia's domestic low cost carrier passenger demand using a genetic algorithm approach. *Aviation*, *20*(2), 39-47. https://doi.org/10.3846/16487788.2016.1171798

[23] Scholz-Reiter, B., Kück, M., & Lappe, D. (2014). Prediction of customer demands for production planning–Automated selection and configuration of suitable prediction methods. *CIRP annals*, *63*(1), 417-420. https://doi.org/10.1016/j.cirp.2014.03.106

[24] Carbonneau, R., Vahidov, R., & Laframboise, K. (2007). Machine learning-Based Demand forecasting in supply chains. *International journal of intelligent information technologies (IJIIT)*, *3*(4), 40-57. https://doi.org/10.4018/jiit.2007100103

[25] Vairagade, N., Logofatu, D., Leon, F., & Muharemi, F. (2019, September). Demand forecasting using random forest and artificial neural network for supply chain management. *International conference on computational collective intelligence* (pp. 328-339). Springer, Cham. https://doi.org/10.1007/978-3-030-28377-3_27

[26] Ampazis, N. (2015). Forecasting demand in supply chain using machine learning algorithms. *International journal of artificial life research (IJALR)*, *5*(1), 56-73. https://doi.org/10.4018/IJALR.2015010104

[27] Efendigil, T., Önüt, S., & Kahraman, C. (2009). A decision support system for demand forecasting with artificial neural networks and neuro-fuzzy models: a comparative analysis. *Expert systems with applications*, *36*(3), 6697-6707. https://doi.org/10.1016/j.eswa.2008.08.058

[28] Chen, Y., Zhao, H., & Yu, L. (2010). Demand forecasting in automotive aftermarket based on ARMA model. *2010 international conference on management and service science* (pp. 1-4). IEEE. https://doi.org/10.1109/ICMSS.2010.5577867

[29] Carbonneau, R., Laframboise, K., & Vahidov, R. (2008). Application of machine learning techniques for supply chain demand forecasting. *European journal of operational research*, *184*(3), 1140-1154. https://doi.org/10.1016/j.ejor.2006.12.004

[30] Zhang, X. (2004). Evolution of ARMA demand in supply chains. *Manufacturing & Service operations management*, *6*(2), 195-198. https://doi.org/10.1287/msom.1040.0042

[31] Du, X. F., Leung, S. C., Zhang, J. L., & Lai, K. K. (2013). Demand forecasting of perishable farm products using support vector machine. *International journal of systems science*, *44*(3), 556-567. https://doi.org/10.1080/00207721.2011.617888

[32] Chan, H. K., Xu, S., & Qi, X. (2019). A comparison of time series methods for forecasting container throughput. *International journal of logistics research and applications*, *22*(3), 294-303. https://doi.org/10.1080/13675567.2018.1525342

[33] Yahya, N. A., Samsudin, R., Shabri, A., & Saeed, F. (2019). Combined group method of data handling models using artificial bee colony algorithm in time series forecasting. *Procedia computer science*, *163*, 319-329. https://doi.org/10.1016/j.procs.2019.12.114

[34] de Oliveira, J. F. L., Pacífico, L. D. S., de Mattos Neto, P. S. G., Barreiros, E. F. S., de Oliveira Rodrigues, C. M., & de Almeida Filho, A. T. (2020). A hybrid optimized error correction system for time series forecasting. *Applied soft computing*, *87*, 105970. https://doi.org/10.1016/j.asoc.2019.105970

[35] Sezer, O. B., Gudelek, M. U., & Ozbayoglu, A. M. (2020). Financial time series forecasting with deep learning: a systematic literature review: 2005–2019. *Applied soft computing*, *90*, 106181. https://doi.org/10.1016/j.asoc.2020.106181

[36] Adeniran, A. O., Kanyio, O. A., & Owoeye, A. S. (2018). Forecasting methods for domestic air passenger demand in Nigeria. *Journal of applied research on industrial engineering*, 5(2), 146–155 http://repository.futminna.edu.ng:8080/jspui/handle/123456789/8743

[37] Minli, Z., & Shanshan, Q. (2012). Research on the application of artificial neural networks in tender offer for construction projects. *Physics procedia*, *24*, 1781-1788. https://doi.org/10.1016/j.phpro.2012.02.262

[38] Movahedi Sobhani, F., & Madadi, T. (2015). Studying the suitability of different data mining methods for delay analysis in construction projects. *Journal of applied research on industrial engineering*, *2*(1), 15-33.

[39] Nury, A. H., Hasan, K., & Alam, M. J. B. (2017). Comparative study of wavelet-ARIMA and wavelet-ANN models for temperature time series data in northeastern Bangladesh. *Journal of King Saud university-science*, *29*(1), 47-61. https://doi.org/10.1016/j.jksus.2015.12.002

[40] Zhang, G., Patuwo, B. E., & Hu, M. Y. (1998). Forecasting with artificial neural networks: the state of the art. *International journal of forecasting*, *14*(1), 35-62. https://doi.org/10.1016/S0169-2070(97)00044-7

[41] Li, G. Q., Xu, S. W., & Li, Z. M. (2010). Short-term price forecasting for agro-products using artificial neural networks. *Agriculture and agricultural science procedia*, *1*, 278-287. https://doi.org/10.1016/j.aaspro.201 0.09.035

[42] Moreno, J. J. M. (2011). Artificial neural networks applied to forecasting time series. *Psicothema*, *23*(2), 322-329.

[43] Olawoyin, A., & Chen, Y. (2018). Predicting the future with artificial neural network. *Procedia computer science*, *140*, 383-392. https://doi.org/10.1016/j.procs.2018.10.300

[44] Lo, D. C., Wei, C. C., & Tsai, E. P. (2015). Parameter automatic calibration approach for neural-network-based cyclonic precipitation forecast models. *Water*, *7*(7), 3963-3977. https://doi.org/10.3390/w7073963

[45] Faris, H., Aljarah, I., & Mirjalili, S. (2017). Evolving radial basis function networks using moth–flame optimizer. In *Handbook of neural computation* (pp. 537-550). Academic Press. https://doi.org/10.1016/B978-0-12-811318-9.00028-4

[46] Tatar, A., Naseri, S., Sirach, N., Lee, M., & Bahadori, A. (2015). Prediction of reservoir brine properties using radial basis function (RBF) neural network. *Petroleum*, *1*(4), 349-357. https://doi.org/10.1016/j.petlm.2015.10.011

[47] Shabri, A., & Samsudin, R. (2014). A hybrid GMDH and box-jenkins models in time series forecasting. *Applied mathematical sciences*, *8*(62), 3051-3062. https://doi.org/10.12988/am s.2014.44270

[48] Ivakhnenko, A. G. (1968). The group method of data handling, a rival of the method of stochastic approximation. *Soviet automatic control*, *13*(3), 43-55.

[49] Farlow, S. J. (1981). The GMDH algorithm of Ivakhnenko. *The American statistician*, *35*(4), 210-215. https://www.researchgate.net/publication/254330626_The_GMDH_algorithm_of_Ivakhnenko

[50] Anastasakis, L., & Mort, N. (2001). *The development of self-organization techniques in modelling: a review of the group method of data handling (GMDH)* (Report No. 813). University of Sheffield, United Kingdom. https://gmdhsoftware.com/GMDH_%20Anastasakis_and_Mort_2001.pdf

[51] Ivakhnenko, A. G., & Müller, J. A. (1995). Self-organisation of nets of active neurons. *System analysis modeling simulation*, *20*(1–2), 93–106.

[52] Ziari, H., Sobhani, J., Ayoubinejad, J., & Hartmann, T. (2016). Analysing the accuracy of pavement performance models in the short and long terms: GMDH and ANFIS methods. *Road materials and pavement design*, *17*(3), 619-637. https://doi.org/10.1080/14680629.2015.1108218

[53] Jang, J. S. (1993). ANFIS: adaptive-network-based fuzzy inference system. *IEEE transactions on systems, man, and cybernetics*, *23*(3), 665-685. DOI: 10.1109/21.256541

[54] Takagi, T., & Sugeno, M. (1983). Derivation of fuzzy control rules from human operator's control actions. *IFAC proceedings volumes*, *16*(13), 55-60. https://doi.org/10.1016/S1474-6670(17)62005-6

[55] Aengchuan, P., & Phruksaphanrat, B. (2018). Comparison of fuzzy inference system (FIS), FIS with artificial neural networks (FIS+ ANN) and FIS with adaptive neuro-fuzzy inference system (FIS+ ANFIS) for inventory control. *Journal of intelligent manufacturing*, *29*(4), 905-923. https://doi.org/10.1007/s10845-015-1146-1

[56] Sremac, S., Kazimieras Zavadskas, E., Matić, B., Kopić, M., & Stević, Ž. (2019). Neuro-fuzzy inference systems approach to decision support system for economic order quantity. *Economic research-ekonomska istraživanja*, *32*(1), 1114-1137. https://doi.org/10.1080/1331677X.2019.1613249

[57] Alexandridis, A. K., & Zapranis, A. D. (2013). Wavelet neural networks: a practical guide. *Neural networks*, *42*, 1-27. https://doi.org/10.1016/j.neunet.2013.01.008

[58] Gürsoy, Ö., & Engin, S. N. (2019). A wavelet neural network approach to predict daily river discharge using meteorological data. *Measurement and control*, *52*(5-6), 599-607. https://doi.org/10.1177/0020294019827972

[59] Kærgaard, K., Jensen, S. H., & Puthusserypady, S. (2016). A comprehensive performance analysis of EEMD-BLMS and DWT-NN hybrid algorithms for ECG denoising. *Biomedical signal processing and control*, *25*, 178-187. https://doi.org/10.1016/j.bspc.2015.11.012

[60] Khosravi, A., Machado, L., & Nunes, R. O. (2018). Time-series prediction of wind speed using machine learning algorithms: a case study Osorio wind farm, Brazil. *Applied energy*, *224*, 550-566. https://doi.org/10.1016/j.apenergy.2018.05.043

[61] Sun, G., Jiang, C., Cheng, P., Liu, Y., Wang, X., Fu, Y., & He, Y. (2018). Short-term wind power forecasts by a synthetical similar time series data mining method. *Renewable energy*, *115*, 575-584. https://doi.org/10.1016/j.renene.2017.08.071

[62] Al-hnaity, B., & Abbod, M. (2016). Predicting financial time series data using hybrid model. In *Intelligent systems and applications* (pp. 19-41). Springer, Cham. https://doi.org/10.1007/978-3-319-33386-1_2

[63] Neill, S. P., & Hashemi, M. R. (2018). Ocean modelling for resource characterization. *Fundamentals of ocean renewable energy*, 193-235. https://doi.org/10.1016/B978-0-12-810448-4.00008-2