ENGINEERING RESEARCH JOURNAL (ERJ)
Vol. 1, No. 50 October 2021, pp.29-40
Journal Homepage: erj.bu.edu.eg

Comparative Study between Classical Methods (CM) and Machine Learning Algorithms (MLA) for Time Series Forecasting

Heba Salah*, Mohammed Hussein, and Ismail Zahran*****

*** Master Eng. Student, ** Prof of IE, Helwan University, *** Assist. Prof, Helwan University**

Abstract: Accurately demand forecasting is vital for effective supply chain planning activities and all related applications. The demand patterns which generated to forecast including horizontal pattern which happens when the data values vary horizontally around a constant mean. For example, a product whose sales are stable and do not increase or reduce such as dairy milk. Seasonal pattern occurs if a series depending on seasonal factors (e.g., the quarter of the year, the month, or the day of the week such as flu vaccine, soft drinks and ice creams. A trend pattern When data shows a long-term growth or decline. An example of a trend pattern in many firms' sales, businesses, such as electrical vehicles and cell phones. However, all forecasting models have distinct advantages and limitations where the generally accepted principle that no individual forecasting model is the best for all situations under all circumstances. Selecting appropriate forecasting methods from numerous alternatives is crucial to success. In this work, we conduct a comparative study between classical and machine learning forecasting algorithms via a statistical programming language R, which is used to generate time series data. The generated data has a mean of 2000 units and standard deviation ranges from 10 to 50, and has a different factor that influence the forecasting ability of classical methods and machine learning algorithms in how to utilize their capacity and extract information effectively. As the amount of historical data available, the type of data pattern, increasing or decreasing trend or seasonal factor, and the variation amount or randomness available in the data. The performance was measured using (MAPE) as the accuracy measurement of demand forecasting.

Keywords: **:(** Demand Forecasting, Neural Networks, Timeseries, MAPE)

1. Introduction

Time series forecasting is important and essential for all planning and decision-making tasks. It builds a competitive edge in today's competitive market. It's applications extent from inventory management and scheduling to planning strategies in numerous fields such as economics, meteorology, sales, industry, science, and engineering. Many forecasting techniques are available to manage from the highly complex approaches such as machine learning to the simplest naive model. Selection of a forecasting method depends on the characteristics of data available to make forecasting process such as mean, standard deviation, type of data pattern, availability of historical data or run length, and the cost or the benefit of the forecast to the company. The study results were used to establish an aggregated protocol to select the appropriate forecasting methodology based on the forecasting data

Table.1 below summarizes the classical forecasting approaches used in this research

Table1: Classical Forecasting Approaches

MODEL	*Forecasting method*	*Definition of forecasting method*
CLASSICA	Naïve Forecast	• The method which uses the latest value of the variable as the future forecast. $F_{t+1}=Y_t$ (1) where F_{t+1} is a forecast of the times series for period (t + 1) , Y_t is an actual value of the time series in period (t)
	MA	• This method forecasts the next period using the average of the most recent (k) data values in the time series. $F_{t+1}=\frac{1}{k}\sum_{i=t-k+1}^{t} Yi$ (2)where F_{t+1} is a forecast of the times series for period (t + 1), Y_t is an actual value of the time series in period (t) .
	DMA	• The centered moving averages are an example of how a moving average can itself be smoothed by another moving average. F_{t+1}= 1/k $\sum_{i=1}^{k} Ti$ (3) ,where $T_{i\ is}$ MA of previous periods, i=1, 2,..k.
	EXP	• This method selects only one weight which is the weight for the most recent observation and other values calculated automatically. $F_0 = Y_0$, $F_{t+1} = \alpha Y_t + (1 - \alpha) F_t = F_t + \alpha (Y_t - F_t)$ (4) • where α ∈ [0, 1] is the smoothing factor, F_{t+1} is the new forecast for next period, Y_t is the current observation at period t, and Ft is the last forecast made in period t−1
	ADAP	$F_{t+1}= \alpha_t Y_t+(1- \alpha_t) F_t$ (5) Where $\alpha_{t+1} = \| A_t / M_t \|$ (6) $A_t=\beta E_t+(1- \beta) A_{t-1}$ (7) $M_t=\beta \|E_t\|+(1-\beta) M_{t-1}$ (8) $E_t=Y_t - F_t$ (9) Where | | denotes absolute values, and B is a parameter between 0 and 1.A smoothed calculation of forecast error is estimated as a weighted average of A_{t-1} and the last forecasting error E_t. M_t is a smoothed approximation of the absolute forecast error, measured as a weighted average of M_{t-1} and the most recent absolute forecasting error $\|E_t\|$.
	H	• It is using for two smoothing constants one to account for the level of the time series and the other for linear trend in the data .$L_t= \alpha Y_t +(1- \alpha) (L_{t-1}+b_{t-1})$ (10) $B_t= \beta(Lt-L_{t-1}) +(1- \beta) b_{t-1}$ (11) $F_{t+m}=Lt +b_t m$ (12) The two smoothing constants, α and β (with values between 0 and 1), • L_t denotes an approximation of the series' level at time t, while b_t denotes an estimate of the series' slope at time t
	H&W	• The Holt-Winters' method is based on three smoothing Level: $L_t = \alpha Y_t / S_{t-s} +(1- \alpha) (L_{t-1} + b_{t-1})$ (13) Trend: $b_t = \beta (L_t-L_{t-1}) +(1- \beta) b_{t-1}$ (14) Seasonal: $S_t = \gamma Y_t / L_t + (1- \gamma) S_{t-s}$ (15) Forecast: $F_{t+m} = (L_t + b_t m) S_{t-s+m}$ (16)

1.2 Machine Learning Algorithms (MLA)

A neural network is a network (ANN) made up of neurons that are arranged in layers. The predictors are included in the first layer (input layer), the last layer (output layer) and there are intermediate layers in between (hidden layers). The nodes play an important role in the network where it constructs and transforms of inputs into outputs as shown in figure1. The weights are used as coefficients in the prediction process to amplify or minimize the input signal to each neuron in the network. Each layer receives inputs from the previous layer, which performs a nonlinear transformation on them before passing them on to the next layer. The network's processing ability is stored in the inter-unit connection weights, which are learned from a series of training patterns. Learning is the process of changing the connected weights between nodes as a result of the difference between the targeted output and the expected output of a neural network in response to input data at the input layer. There're many neural network algorithms such as, Backpropagation algorithm which used in this study.

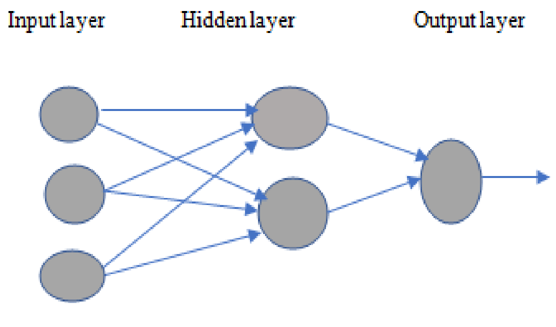

Fig 1: One Hidden Layer -NN

1.2.1 Back-Propagation Neural Networks (BPNN)

The choosing the best BPNN design depend on the minimum deviation between the targeted and expected value of the output. The selection of the learning rate, momentum rate, and the number of hidden layers need of many trials and error. The Learning Rate (0-1) helps to converge the NN training process while the Momentum Rate (0-1) helps to accelerate the training process. The working mechanism of backpropagation ANN can be explained in the following steps as shown in figure 2:

Step 1. Entered the Input and output vectors into the system.

Step 2. Network assigns the parameters of to the inputs randomly.

Step 3. Calculate the error (differences between the expected and real outputs).

Step 4. Minimize errors by adjusting parameters in the appropriate direction.

Step 5. Repeating The learning process until the network encounters an error.

The research work in forecasting is classified by the scope of application: -

1.3 Demand forecasting scope in supply chain management

In this field the researchers concerned to obtain an accurate prediction of future demand supply chain to decrease bullwhip effect in supply chain which caused due to the amplification of the process of order in upward direction**. Ma, J., & Bao, B. (2017);** **[1]** discussed the factors effected in selection of forecasting method such as size of historical data and standard deviation. **Alzubaidi, Z. Y. (2020);** **[2]** presented a comparative study between classical forecasting methods and machine learning algorithms to investigate the optimal method to forecast shipment data of an FMCG company. On other cases the researchers tried to use more than one forecasting method or scenarios to overcome the supply chain forecasting problems for example, in pharmaceutical supply chain, **Merkuryevaa, Aija Valbergab, Galina, Alexander Smirnov. (2018);** **[3]** introduced three forecasting scenarios: SMA model, multiple linear regressions, and symbolic regression with genetic programming and **Chirag Deba, Fan Zhangb, Junjing Yanga, Siew Eang Leea, Kwok Wei Shaha.(2018);**][4] summarized the art forecasting methods, its applications and the main factors of selecting a forecasting method. **Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., & Vasilakis, C. (2021).** [5] used and compared statistical, deep learning models, and new hybrid forecasting methods based on nearest neighbors and clustering, for short, mid, and long-term forecasts for excess demand products and services which had a high effect in solving many forecasting problems.

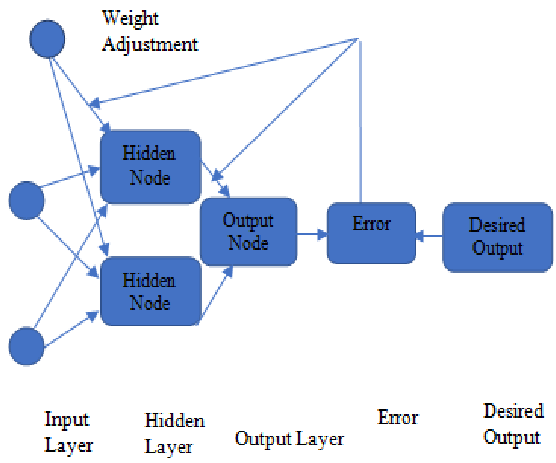

Fig2. General Architecture of Backpropagation NN

1.4 Sales Forecasting: -

In this field the method of predicting future revenue is known as sales forecasting. Companies may make informed business decisions and predict short- and long-term results with accurate revenue forecasts.

Bohdan M. Pavlyshenko. (2018),,[6] compared many time series methods with machine learning and investigated the factors that effect in increasing accuracy for sales predictive analytics. **Al-Maqaleh, B. M., Al-Mansoub, A. A., & Al-Badani, F. N. (2016).** **[7]** compared traditional forecasting methods with machine learning methods and showed the ability to deal with the nonlinear and stochastic data that backpropagation neural network with adaptive slope and momentum parameter gives better forecast for the indices of a consumer in the Republic of Yemen due to their ability to deal with the nonlinear and stochastic data better than Box-Jenkins model. **Nor, M. E., Safuan, H. M., Shab, N. F. M., Asrul, M., Abdullah, A., Mohamad, N. A. I., & Lee, M. H. (2017, May).** **[8]** presented a comparative study between Box-Jenkins models and neural networks algorithms to forecast a gold price data. **Qian, X. Y., & Gao, S. (2017). [9].** compared between time series models and machine learning algorithms by using real data sets to present financial series prediction. **Ghita Benboubker, Dr.Ilham Kissani, Dr. Asmaa Mourhir, 2019; [10]** compared classical methods and neural networks, to predict the sales of the year 2019 via the statistical programming language R and concluded that no single forecasting technique dominated for certain demand pattern or the best for all data sets situations. also, **Aras, S., Deveci Kocakoç, İ., & Polat, C. (2017), [11].** Conducted a comparative study between different models such as the ETS, ANN, ARIMA and the ARFIMA, to forecasted retail sales.

1.5 Demand forecasting scope in machine failures and spare parts

In this field the preparation of the maintenance activities and spare parts is necessary to be forecasted to avoid the high cost due to the failures in machines and the excess inventory. **Ramos, P., Oliveira, J. M. S., & Silva, P. (2014).,[12]** presented a predictive study that was used on manufacturing equipment to predict malfunctions and, compared the forecasting performance of ARIMA and NN models to detect faults and maintenance behavior based on the sensors' future values forecasts. The results showed that the ARIMA model accurately predicts the increase in distance between the discs before and after replacement than NN model. **González Vargas* and M. Elizondo Cortés. (2017),,[13]** Evaluated the monthly demand forecasting of automobile spare parts in Mexico by comparing Autoregressive integrated models of moving averages (ARIMA), artificial neural networks (ANNs), moving averages, exponential smoothing and the ARIMA-ANNs hybrid models. **Bhuvana Adur Kannan1, Ganesh Kodi1, Oscar Padilla1, Doug Gray1,2, and Barry C Smith. (2020),[14]** compared between traditional forecasting techniques like ARIMA, Croston, enhanced Croston (TSB), Facebook Prophet with machine learning models such as: Random Forest, Artificial Neural Networks (Keras RNN), Extreme Gradient Boosting (XG Boost) **Mahindra, Hassan ,Hamzan and Zulkifly;(2018); [15]** compared forecasting techniques such as naïve, exponential smoothing, linear regression, and moving average with neural network and established a collaborative forecasting from external resources to reduce the error and obtaining tacit knowledge in the purchasing process.

1.6 Demand Forecasting Scope in Inventory Management: -

In this field the choosing of the right forecasting method is avital task when making optimization process to companies that wish to reduce the costs attached to their inventory while making sure that they can meet the demand for their products. **Nimai Chand Das Adhikari∗, Nishanth Domakonda**

Chinmaya Chandan∗ Gaurav Gupta∗ , Rajat Garg∗ , S Teja∗ , Lalit Das∗ , Dr. Ashutosh Misra.(2017),[16], compared between different timeseries techniques used in demand forecasting fields such as Weighted moving Average, Simple moving average, Autoregressive integrated moving average (ARIMA) and Holts Winter with regression-based models such as Support Vector Regression, Decision Trees Regression, Linear Regression, Ridge Regression and Random Forest Regression by using R-Studio to generate the time-series model, with an ensemble forecast results . **H., & Rajendran, S. (2019)., [17]** compared between time series forecasting methods with machine learning algorithms of Taiwan blood services foundations blood supply to determine the inventory policy based on the estimated future blood supply. They concluded that there is not a single method that forecasts the supply accurately, and recommend using the average value of the forecasts obtained from different methods.

1.7 Demand Forecasting Scope in Electrical/Energy Demand: -

Divina, F., Garcia Torres, M., Goméz Vela, F. A., & Vazquez Noguera, J. L. (2019); [18] presented a comparative study of time series forecasting methods for short term electric energy consumption for predicting one-day electric energy consumption in non-residential buildings. **Athiyarath, S., Paul, M., & Krishnaswamy, S. (2020).][19]** evaluated many forecasting models of three different time series (time-space) and patterns short, medium, and long-period forecasting.

1.8 Demand Forecasting Scope in Planning

Adeyinka, D. A., & Muhajarine, N. (2020).; [20] studied time series prediction of under-five mortality rates for Nigeria which is essential for policy actions and planning by conducting a comparative analysis of artificial neural networks, Holt-Winters's exponential smoothing and autoregressive integrated moving average models, and holt-winters smoothing exponential methods with GMDH-type neural network (group method of data handling (GMDH)-type artificial neural). **Cecaj, A., Lippi, M., Mamei, M., & Zambonelli, F. (2020) . [21]** Presented a comparative study between Deep Learning and Statistical Methods in Forecasting Crowd Distribution from Aggregated Mobile Phone Data.). **Li, Y., & Cao, H. (2018). [22]** studied prediction for tourism flow by conducting comparative study between machine learning algorithms such as backpropagation neural network and classical methods such as Auto-

Regressive Integrated Moving Average (ARIMA) model

2 Experimental Work

2.1 Generated Data Patterns

The sequence of generated timeseries demand data was taken at successive equally spaced points in time by using R language. Four types of time series patterns can be distinguished: horizontal, seasonal, trend, and trend with seasonality as shown in Table2.

2.1.1 Data Verification

The generated data was verified through plotting and calculating descriptive statistics to make sure about its characteristics.

2.2 Applying Forecasting Models

Table3 summarize description for classical and machine learning forecasting methods and parameter levels used to conduct comparative study between forecasting methods.

2.3 Evaluating a forecasting model

Forecasting models are validated using Mean Absolute Percentage Error (MAPE) to measure the goodness for both classical and machine learning approach as a forecasting performance measure and to compare forecasting models. MAPE is computed as the average of the absolute difference between the forecasted and actual values, expressed as a percentage of the actual values [23]. MAPE indicates how much error in predicting compared with the real value. MAPE= $\sum (|Y_t - F_t / Y_t| \times 100)/n$ [24]

2.4 The Comparative Study Protocol

Figure 3. shows the flow chart diagram of forecasting modeling process to select the proper forecasting method for demand pattern through conducting a comparative study between classical methods and BPNNT algorithms.

3 Results and Discussions

In this section, the results obtained by forecasting the different techniques described in Table 3 on the datasets presented in Table 2 to select the most proper forecasting method for demand pattern and draw the main conclusions.

3.1 Horizontal Data Patterns.

3.1.1 Classical methods

3.1.1.1 Selecting parameter levels for classical forecasting methods.

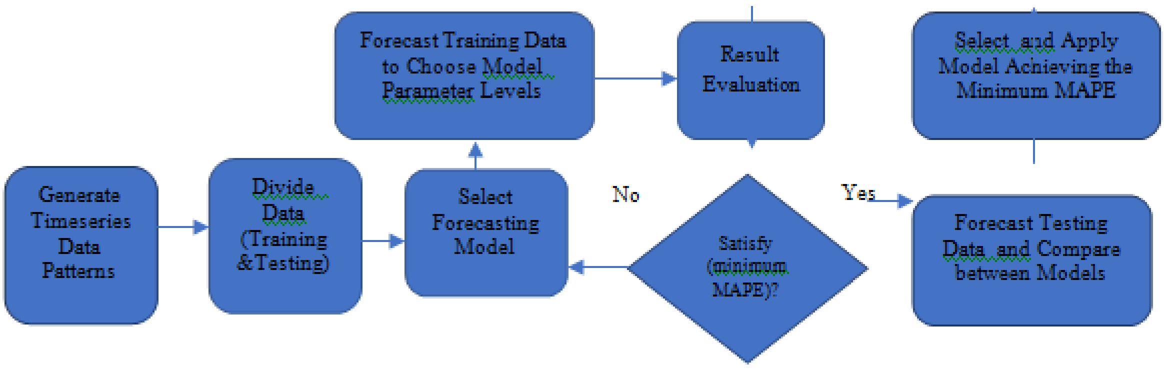

Fig 3: Flowchart diagram of forecasting modeling process

Table 4. shows the parameter levels for classical methods which achieved the minimum level of MAPE of forecasting training data.

4. Selecting Forecasting Method

Table 5. shows the most proper classical methods which achieved the minimum level of MAPE of forecasting testing data.

From Table 5, one can conclude:
1-Vars in the data effected negatively in accuracy of forecasting by classical methods.
2-classical methods proper for forecasting the horizontal data which has divided to (70% training: 30 % testing).
5.1.2. Backpropagation Neural Network Algorithm (BPNNT).

Table 6. shows the BPNNT architecture design for forecasting horizontal demand patterns.

From Table 6, it can be concluded that decreasing accuracy of BPNNT by increasing the Var in data and it is proper to forecast data with dividing percentage (85% training: 15% testing).

Table2: Timeseries Data Sets Generation Scheme

Pattern			Variation (Var) Low	High
Stationary (Horizontal) [Mean = 2000 Units]			Standard Deviation = 10 Units	Standard Deviation = 50 Units
Trend [Intercept = 1000 Units]		Low Slope = 10 Units per Period	Standard Deviation = 10 Units	Standard Deviation = 50 Units
		High Slope = 50 Units per Period	Standard Deviation = 10 Units	Standard Deviation = 50 Units
Stationery and Seasonality [Mean = 2000 Units]		Seasonality Period = 1	Standard Deviation = 10 Units	Standard Deviation = 50 Units
		Seasonality Periods = 4	Standard Deviation = 10 Units	Standard Deviation = 50 Units
Trend and Seasonality [Intercept = 1000 Units]	Low Slope = 10 Units per Period	Seasonality Period = 1	Standard Deviation = 10 Units	Standard Deviation = 50 Units
	High Slope = 50 Units per Period	Seasonality Periods = 4	Standard Deviation = 10 Units	Standard Deviation = 50 Units
	Low Slope = 10 Units per Period	Seasonality Period = 1	Standard Deviation = 10 Units	Standard Deviation = 50 Units
	High Slope = 50 Units per Period	Seasonality Periods = 4	Standard Deviation = 10 Units	Standard Deviation = 50 Units

Table 3: Forecasting Methods Description

Forecasting Methods		Parameters Levels
Classical Methods (CM)	Naïve	0
	Moving Average	k=3,5,7
	Double Moving Average	k=3,5,7
	Exponential Smoothing	Alfa(α)=0.05,0.08,0.2,0.5,0.8
	Adaptive Exponential Smoothing	(Alfa(α), Beta(β)) = (0.05,0.2), (0.08,0.2), (0.2,0.2),0.5,0.2), (0.8,0.2), (0.4,0.05), (0.2,0.08), (0.2,0.5), (0.2,0.8)

Table 4: Parameter Levels of Classical Methods (CM)

CM Forecasting Method	Model Parameter	MAPE
Naive	(No parameter)	0.62
Moving Average (order k)	K=5	0.48
Double Moving Average	K=7	0.47
Exponential Smoothing	α=0.2	0.45
Adaptive Moving Average	α=0.8, β=0.2	0.47
Holt Linear	α=0.5, β=0.2	0.59
Holt Winter	α=0.05, β=0.2	0.67

Table 5: Classical Methods for Horizontal Demand Patterns

Demand Pattern	Forecasting Method	MAPE
Low Var-84	MA=5	0.487
Low Var-168	EXP=0.2	0.53
Low Var-204	DMA=7	0.4508
High Var -84	EXP=0.2	2.531
High Var -168	EXP=0.2	2.637
High Var -204	EXP=0.2	2.275

Table 6: BPNNT Architecture Designs for Forecasting Horizontal Demand Patterns

Demand Pattern	Low Variation-84	Low Variation-168	Low Variation-204	High Variation -84	High Variation -168
Hidden Node	3	15	7	3	5
Input Node	2	2	3	2	2
Learning Rate	0.01	0.01	0.01	0.01	0.001
Momentum	0.3	0.000001	0.000001	0.3	0.000001
Epoch	1000	1000	500	1000	15000
MAPE	0.538	0.47027	0.4934	2.7897	2.408

Table 7: Comparative Study between CM and BPNNT in Forecasting Horizontal Patterns

Forecasting Method	MAPE (BPNNT Algorithms)	MAPE CM	Demand Pattern
MA=5	0.538	0.487	Low Var-84
BPNNT (Hidden Node=15)	0.47027	0.53	Low Var-168
DMA=7	0.4934	0.4508	Low Var-204
EXP=0.2	2.7897	2.531	High Var -84
BPNNT (Hidden Node=5)	2.408	2.637	High Var -168
EXP=0.2	2.665	2.275	High Var -204

Table 8: Comparison between CM and BPNNT for Seasonal Data

Demand Pattern	MAPE (CM)	MAPE (BPNNT)	Forecasting Method
Low Var -Long Season Period-84	0.8835	2.3012	H&W 0.08,0.2
Low Var -Long Season Period-168	0.8748	2.4911	H&W 0.08,0.2
Low Var - Long Season Period-204	0.793	2.0543	H&W 0.08,0.2
High Var -Long Season Period-84	3.97	3.687	BPNNT (Hidden Node =45)
High Var -Long Season Period-168	3.432	4.417	H&W 0.05,0.2
High Var -Long Season Period-204	3.185	2.9628	BPNNT (Hidden Node=7)
Low Var -Short Season Period-84	0.738	1	H&W= 0.05,0.2
Low Var -Short Season Period-168	0.637	1.372	H&W= 0.05,0.2
Low Var -Short Season Period-204	0.691	1.1944	H&W= 0.05,0.2
High Var -Short Season Period-84	3.081	2.2881	BPNNT (Hidden Node=1)
High Var -Short Season Period-168	3.046	6.872	H&W= 0.05,0.2
High Var -Short Season Period-204	3.066	2.8557	BPNNT (Hidden Node=10)

4.2 Data Patterns with Trend.

Table 9. shows the comparative MAPE between classical methods and BPNNT algorithms in forecasting data patterns with trend.

In case of data patterns with trend,

1- Classical methods outperform BPNNT in forecasting trend data patterns where holt linear exponential smoothing method success in forecasting most of trend data.
2- Increasing run length of data plays a prominent role in increasing forecasting accuracy of classical methods
3- BPNNT has a good ability to forecast data with dividing percentage (85% training :15% testing).

4- Increasing trend factor and Var in data decreasing the accuracy of forecasting low Var trend data

5 Conclusions

The conclusion of conducting a comparative study between classical methods and BPNNT algorithms are: -

1. If there's Variation in the data but it moving around a certain average it advised to forecast by neural network algorithm, where it is obvious when the neural network algorithm gives better forecasting performance in short run length, horizontal data pattern.
2. Increasing Var in data increase data randomness which has a bad effect in accuracy of forecasting methods

3. There's no specific forecasting method is dominated for certain type of data pattern as general principle state that "No single forecasting technique is the best for all data sets with respect to all statistics considered and no individual forecasting model is the best for all situations under all circumstances", which supported by **(**Aras, S., Deveci Kocakoç, İ., & Polat, C. (2017) [11].

4. As aggregated results of previous study, the classical methods are more convenient to forecast this domain of data than machine learning algorithms because it relies on past data and they are more accurate when we can assume that some of the past trends will extend in the future as agreed with (Hyndman and Athanasopoulos 2018) [23] than the forecasting by neural network which is work as black box.

5. Classical forecasting methods proper for forecasting short-term times series than machine learning algorithm which agreed with the results obtained from study conducted by Makridakis et al. (2018) [24], which was obvious in forecasting horizontal and trend pattern and other patterns from seasonal and trend data.

6. BPNNT forecast effectively high Variation data due to its ability to deal with non-linearity in data which was obvious in:

   A- High Variation, long seasonal pattern, short run length and long run length
   B- High Variation, short seasonal pattern, short run length and long run length
   C- High Variation, horizontal, medium run length pattern
   D- High Variation, short seasonal with high trend pattern, short run length and medium run length

7. Selection forecasting method depends on the characteristics of data, which is obvious in:

   A- Increasing run length of data prominent in increasing accuracy of trend pattern, seasonal with trend pattern, long seasonality pattern.
   B- The percentage of dividing data is critical in forecasting by BPNNT which is worked well for forecasting seasonal pattern and horizontal pattern which has (70% training data:30% testing data). Also, classical methods are proper for forecasting seasonal, horizontal and trend patterns which have data division equal (85% training data:15% testing data).
   C- Increasing trend factor and seasonal period is prominent in increasing the accuracy of trend pattern and trend with seasonal pattern.
   D- Increasing the length of seasonal period is prominent in increasing the accuracy of seasonal pattern and seasonal with trend demand pattern.

8. Establish an aggregated forecasting method protocol

Table 9: Comparison between CM and BPNNT for Data with Trend

Demand Patterns	MAPE CM	MAPE BPNNT Algorithms	Forecasting Methods
low Var-low trend-84	0.6925	2.221	H=0.5,0.2
low Var-low trend-168	0.5893	1.2567	H=0.5,0.2
low Var-low trend-204	0.4438	3.56	H=0.5,0.2
high Var-low trend-84	3.061	3.787	MA=3
high Var-low trend-168	2.079	2.394	MA=3
high Var-low trend-204	1.0899	4.59188	MA=3
low Var - high trend-84	0.2873	2.727	H=0.8,0.2
low Var - high trend-168	0.1651	1.533	H=0.8,0.2
low Var- high trend-204	0.1331	4.628	H=0.8,0.2
high Var-high trend-84	1.437	4.187	H=0.8,0.2
high Var-high trend-168	0.8267	1.813	H=0.8,0.2
high Var-high trend-204	0.6666	4.998	H=0.8,0.2

Table 10: Comparison between CM and BPNNT for Data with Horizontal pattern

2-Protocol for Forecasting Horizontal Time Series Patterns				
Horizon	Var		FORECASTING METHOD-A	Forecasting Method-B
Level Run-Length	LOW	HIGH		
short	×		Moving average(k=5)	BPNNT(HD=3)
		×	Exponential smoothing (0.2)	BPNNT(HD=3)
medium	×		BPNNT(Hidden-node=5)	Exponential smoothing (0.2)
		×	Exponential smoothing (0.2)	BPNNT(HD=7)
long	×		BPNNT(Hidden-node=15)	DMA=7
		×	Exponential smoothing (0.2)	BPNNT(HD=7)

Table 11: Comparison between CM and BPNNT for Data with Trend pattern

3-Protocol for Forecasting Trend Time Series Patterns						
Run-Length	Var		Trend		Forecasting Method-A	Forecasting Method-B
	LOW	HIGH	LOW	HIGH		
short	×		×		HOLT (0.5,0.2)	NAÏVE
		×	×		Moving average(k=3)	NULL
	×			×	HOLT (0.8,0.2)	NAÏVE
		×		×	HOLT (0.8,0.2)	NAÏVE
medium	×		×		HOLT (0.5,0.2)	NAÏVE
		×	×		Moving average(k=3)	EXP=0.5
	×			×	HOLT (0.8,0.2)	NAÏVE
		×		×	HOLT (0.8,0.2)	NAÏVE
long	×		×		HOLT (0.5,0.2)	NAÏVE
		×	×		Moving average(k=3)	EXP=0.5
	×			×	HOLT (0.8,0.2)	ADAP=0.8,0.2
		×	×		HOLT (0.8,0.2)	NAÏVE

Table 12: Comparison between CM and BPNNT for Data with Seasonality pattern

4- Protocol for Forecasting seasonality Time Series Patterns						
Run-Length	Var		Seasonal period		Forecasting Method-A	Forecasting Method-B
	LOW	HIGH	LONG	SHORT		
short	×		×		Holt-Winter (0.05,0.2)	NULL
		×	×		Exponential smoothing (0.8)	BPNN(H=45)
	×			×	Holt-Winter (0.05,0.2)	BPNN(H=35)
		×		×	Double moving(k=7)	EXP=0.2
medium	×		×		HOLT-WINTER (0.08,0.2)	NULL
		×	×		Holt-Winter (0.05,0.2)	BPNN(H=45)
	×			×	Holt-Winter (0.05,0.2)	BPBB(H=45)
		×		×	Exponential smoothing (0.08)	H&W=0.5,0.2

Table 12. To be continued

	×		×		Holt-Winter (0.08,0.2)	NULL
		×	×		BPNNT(Hidden-node=1)	NULL
	×			×	Holt-Winter (0.05,0.2)	BPNN(HD=45)
		×		×	BPNNT(Hidden-node=10)	EXP=0.08

Note: NULL Means the used model in plan A is superior to other techniques.

- Abbreviation Definition

Definition	Abbreviation	Definition	Abbreviation
Classical Forecasting Methods	CM	Adaptive-response-rate single exponential smoothing	ADAP
Machine Learning Algorithms	MLA	Holt’s linear method (H)	H
Mean Absolute Percentage Error	MAPE	Holt-Winters' trend and seasonality method	H&W
Variation amount in data	VAR	Artificial Neural Networks	ANN
Exponential smoothing method	EXP	Back-Propagation Neural Networks	BPNNT
Moving Average of Oder K method	MA	Learning Rate	LR
Double Moving Average	DMA	Momentum	MOM

References

[1] Ma, J., & Bao, B. “Research on bullwhip effect in energy-efficient air conditioning supply chain,” Journal of Cleaner Production, vol. 143,pp. 854-865,2017

[2] Alzubaidi, Z. YA “Comparative Study on Statistical and Machine Learning Forecasting Methods for an FMCG Company”. (2020).

[3] Merkuryeva, G., Valberga, A., & Smirnov, A. “Demand forecasting in pharmaceutical supply chains: A case study,” Procedia Computer Science,vol. 149,pp. 3-10, 2019

[4] Dhakry, Nonihal Singh. "Some aspects of policy structuring for strategic planing towards sustainable integrated renewable energy use for rural india." (2018).

[5] Nikolopoulos, K., Punia, S., Schäfers, A., Tsinopoulos, C., & Vasilakis, C.” Forecasting and planning during a pandemic: COVID-19 growth rates, supply chain disruptions, and governmental decisions,” European Journal of Operational Research,vol. 290(1),pp. 99-115 ,2021.

[6] Pao, James J., and Danielle S. Sullivan. "Time Series Sales Forecasting." Final year project, Computer Science, Stanford Univ., Stanford, CA, USA (2017).

[7] Al-Maqaleh, B. M., Al-Mansoub, A. A., & Al-Badani, F. N.” Forecasting using artificial neural network and statistics models. International Journal Education and Management Engineering,”vol. 3,pp. 20-32, 2016

[8] Nor, M. E., Safuan, H. M., Shab, N. F. M., Asrul, M., Abdullah, A., Mohamad, N. A. I., & Lee, M. H. “Neural network versus classical time series forecasting models,” In AIP Conference Proceedings May ,2017, (Vol. 1842, No. 1, p. 030027). AIP Publishing LLC.

[9] Qian, Xin-Yao, and Shan Gao. "Financial series prediction: Comparison between precision of time series models and machine learning methods." arXiv preprint arXiv:1706.00948, pp.1-9, 2017.

[10] Ghita Benboubker, D., Kissani, I., & Mourhir, A. Comparative Analysis in Sales Forecasting: Classical Methods and Neural Networks, Proceedings of the International Conference on Industrial Engineering and Operations Management Toronto, Canada, October 23-25, 2019.

[11] Aras, S., Deveci Kocakoç, İ., & Polat, C. “Comparative study on retail sales forecasting between single and combination methods,” Journal of Business Economics and Management,vol. 18(5),pp. 803-832,2017.

[12] Ramos, P., Oliveira, J. M. S., & Silva, P., Predictive maintenance of production equipment based on neural network autoregression and ARIMA. In 21st International Euro MA Conference-Operations Management in an Innovation Economy, 2014.

[13] Vargas, C. G., & Cortés, M. E. “Automobile spare-parts forecasting: A comparative study of time series methods,”. International Journal of Automotive and Mechanical Engineering, Vol. 14, pp.3898-3912, 2017.

[14] Adur Kannan, B., Kodi, G., Padilla, O., Gray, D., & Smith, B. C., Forecasting Spare Parts Sporadic Demand Using Traditional Methods and Machine Learning-a Comparative Study. SMU Data Science Review, 3(2), 9. (2020).

[15] Nair, M. K. S., Mohamad, H., Jamil, H. A., & Radzi, Z. M.”Forecasting Military Vehicle Spare Parts Requirement using Neural Networks followed by Application of Tacit Knowledge,”. International Journal of Engineering & Technology, vol. 7(4.29),pp. 13-17, 2018.

[16] Adhikari, Nimai Chand Das, Nishanth Domakonda, Chinmaya Chandan, Gaurav Gupta, Rajat Garg, S. Teja, Lalit Das, and Ashutosh Misra. "An Intelligent Approach to Demand T Forecasting mkkS.".

[17] Shih, H., & Rajendran, “Comparison of time series methods and machine learning algorithms for forecasting Taiwan Blood Services Foundation’s blood supply.” Journal of healthcare engineering, 2019.

[18] Divina, F., Garcia Torres, M., Goméz Vela, F. A., & Vazquez Noguera, J. L. “A comparative study of time series forecasting methods for short term electric energy consumption prediction in smart buildings”, Energies, 2019, vol. 12(10), pp. 1934,

[19] Athiyarath, S., Paul, M., & Krishnaswamy, S. “A “Comparative Study and Analysis of Time Series Forecasting Techniques”. SN Computer Science, 2020, vol.1pp. 1-7.

[20] Adeyinka, D. A., & Muhajarine, N. “Time series prediction of under-five mortality rates for Nigeria: comparative analysis of artificial neural networks, Holt-Winters’s exponential smoothing and autoregressiveintegrated moving average models,”. BMC medical research methodology, 2020, vol. 20(1),pp. 1-11.

[21] Cecaj, A., Lippi, M., Mamei, M., & Zambonelli, F. “Comparing Deep Learning and Statistical Methods in Forecasting Crowd Distribution from Aggregated Mobile Phone Data.Applied Sciences,” 2020, vol.10,pp.65-80,

[22] Li, Y., & Cao, H. “Prediction for tourism flow based on LSTM neural network. Procedia Computer Science, “,2018, vol.129, pp.277-283.

[23] Hyndman, Rob J., and George Athanasopoulos. Forecasting: principles and practice. OTexts, 2018.

[24] Makridakis, S., Spiliotis, E., & Assimakopoulos, V. ,Statistical and Machine Learning forecasting methods: Concerns and ways forward.2018vol, 13(3), p.0194889