2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT)
April 09-11. 2021, ISTTS Surabaya, Indonesia

A backpropagation neural network algorithm in agricultural product prices prediction

2021 3rd East Indonesia Conference on Computer and Information Technology (EIConCIT) | 978-1-6654-0514-0/20/$31.00 ©2021 IEEE | DOI: 10.1109/EIConCIT50028.2021.9431897

1st Haviluddin
Faculty of Engineering
Universitas Mulawarman
Samarinda, Indonesia
haviluddin@unmul.ac.id

2nd Taufikurrahman Khosyi
Faculty of Engineering
Universitas Mulawaran
Samarinda, Indonesia
khosyitaufik97@gmail.com

3rd Hario Jati Setyadi
Faculty of Engineering
Universitas Mulawarman
Samarinda, Indonesia
hario.setyadi@gmail.com

4th Aji Prasetya Wibawa
Jurusan Teknik Elektro
Universitas Negeri Malang
Malang, Indonesia
aji.prasetya.ft@um.ac.id

5th Felix Andika Dwiyanto
Graduate School
Universitas Negeri Malang
Malang, Indonesia
felix.adwiyanto@gmail.com

6th Andri Pranolo
Hohai University
Nanjing, China
andri.pranolo@tif.uad.ac.id

7th Fachrul Kurniawan
Magister of Informatic Engineering
University of Islamic State Maulana Malik Ibrahim
Malang, Indonesia
fachrulk@ti.uin-malang.ac.id

8th Muslimin
Faculty of Engineering
Universitas Mulawarman
Samarinda, Indonesia
muslimin@ft.unmul.ac.id

***Abstract*—The price of chili is one of the food commodities that can affect the inflation rate. Its uncertain price and even increasing at certain times will negatively impact society and state. The price offer for chili is still highly dependent on the amount of chili produced. At the same time, the amount of chili productivity depends on the harvested area and land productivity. When the supply in the market is lacking, the price will increase far from its average price. Otherwise, when the supply is excessive, the price will fall far below the regular price. Therefore, it requires a method to estimate this chili's price to support decision making related to price issues. Many forecasting methods have been used to predict data, such as Backpropagation Neural Networks (BPNN) and Single Moving Average (SMA), proven in some cases to provide good forecasting results. These two methods will be compared with the lowest error rate and the best method in predicting chili's price. The results of this research will help various parties as a consideration in making decisions and planning.**

Keywords—prediction, BPNN, SMA, MSE, chili

I. Introduction

The Republic of Indonesia is an archipelagic country with natural resources in agriculture for domestic food and export to other countries [1]. It is a good fact because food is a basic human need that must be fulfilled every day. One of the primary commodities that are widely produced is chili. It is an essential horticultural commodity in Indonesia that is consumed by most of the population [2]. Also, this commodity is very prospective and potential to improve the standard of living of farmers [3]. Market demand for chili products tends to increase from time to time in line with the increase in average consumption in various countries. Hence, the price of chili is getting higher and more fluctuating [4].

Based on that fact, an analysis is needed to model and predict chili's price in the following year to overcome fluctuations [5], [6]. This research object focuses on the Penembolum Senaken market in Tana Paser, East Kalimantan, Indonesia. It is because this market is the central market that provides basic needs, especially chili. For this reason, it is necessary to do forecasting or price prediction of chili. Forecasting or prediction is an activity to predict what will happen in the future [7], [8]. Predictions are usually made to reduce uncertainty about something that will happen in the future [9], [10]. Two models will be compared to determine the best method for predicting chili price. The models are Backpropagation Neural Network (BPNN) method and the Single Moving Average (SMA) method. These methods are the same as analyzing time-series data to determine patterns based on past data and has high accuracy [11], [12]. Other than that, it is excellent in predicting data that the fluctuating movements are stable. The BPNN method is widely used in Artificial Neural Networks, and SMA learning is widely used in statistics.

BPNN is a supervised learning method that has a target to be wanted. Its characteristic is to minimize errors in the output generated by the network. In the BPNN method, a multilayer network is usually [13], [14]. Meanwhile, SMA is a forecasting method that uses several actual new demand data to generate a forecast value for future demand [15]. Furthermore, this paper will apply two models, namely SMA and BPNN that have been developed and compared in order to predict the chili prices. Section 2 describes the architectures of SMA and BPN prediction models are proposed. Section 3 describes the analysis and discussion of the results. Finally, conclusions are summarized in Section 4.

II. Method

The data used in this study are secondary, and the materials used are data of chili in the Penembolum Senaken market. The data analysis technique used in this study was the Backpropagation Neural Network and Single Moving Average methods.

A. Variables

The variable used in this study is the Independent Variable, data of the chili price per day from 2015 to 2018. The presentation of the data shown in Fig 1.

B. Analysis Method

In this study, there are several stages of the analysis method. Those stages are a study of expert knowledge literature, identification of chili prices, data preparation, making predictions using the BPNN and SMA methods, calculating the accuracy of the prediction results from the two methods, the accuracy of the prediction results, the comparison of the accuracy, and the comparison results [16]. The research stages can be seen in Fig. 2.

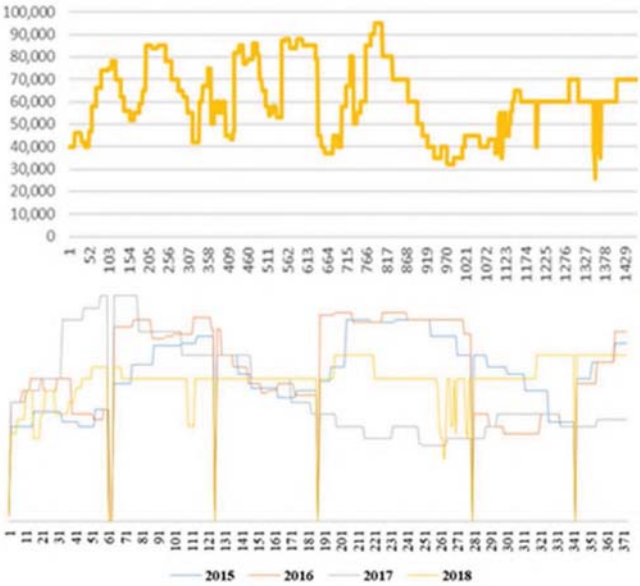

Fig. 1. Price of Chilli 2015-2018

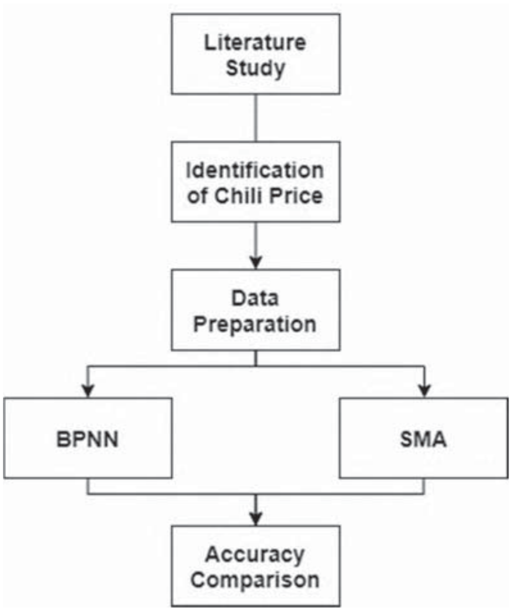

Fig. 2. Research Stages

C. Prediction with BPNN and SMA

It begins to make predictions using the BPNN and SMA methods with existing real data, then calculates the prediction results with these methods. In a BPNN method, the first stage is performing a data normalization. Data of chili per day from 2015 to 2018 is normalized (1).

$$X' = \frac{0.8\,(X-a)x^2}{(b-a)} + 0.1 \tag{1}$$

Then, the second stage of BPNN is data training testing. The training data is obtained from the comprehensive data of 1460 divided by 70%, 80%, and 90%. Data testing was obtained from the overall data as many as 1460 data divided by 30%, 80%, and 10%. Each training testing percentage is 70:30; 80:20; 90:10. Furthermore, the last stage is to test the variables. The data will be normalized, then testing the variables with the parameters that can be seen in Table I to get the best architecture with the smallest MSE value [17].

TABLE I. VARIABLE TEST PARAMETERS

Parameter	
Neuron Hidden Layer	2, 3, 4, 5, 10
Learning Function	trainlm, traingd, traingdx
Learning Rate	0.1;0.2;0.3;0,4;0.5;0,6;0.7;0,8;0.9

Each variable is tested using these parameter variations. The condition that the activation function used in each hidden layer is *tansig*, and the activation function used in the output layer is *purelin*. After each variable is tested with the parameters in Table I, it is selected, and the smallest MSE value is taken. Moreover, the SMA method consists of two stages. The first stage calculates the average price of cayenne pepper per day from 2015 to 2018 with Moving Average 2, 4, 6, and 8 (2). Then it will be continued in a second stage by calculating the error of each forecasting average (3).

$$S_{t+1} = \frac{X_t + X_{t-1} + \dots X_{t-n+1}}{n} \tag{2}$$

$$(\text{rii value} - \text{average of forecasting value})^2 \tag{3}$$

D. Backpropagation Neural Network Method

In forecasting using the BPNN method, several steps can be seen in Fig. 3.

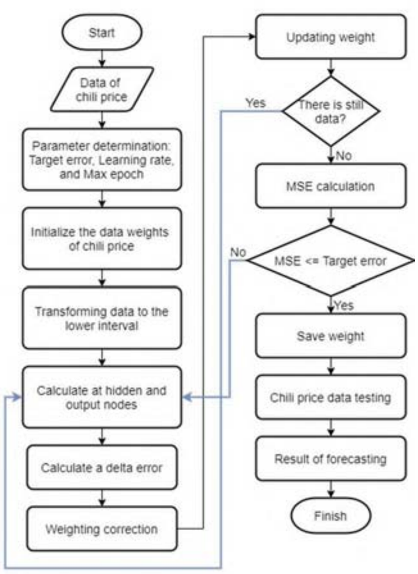

Fig. 3. BPNN Flowchart

The BPNN architecture used to test the variables in this study is presented in Fig. 4, 5, 6, 7 and 8. Each architecture has four input layer neurons according to the independent variable data used, four years, and one neuron output layer according to the dependent variable data, one year.

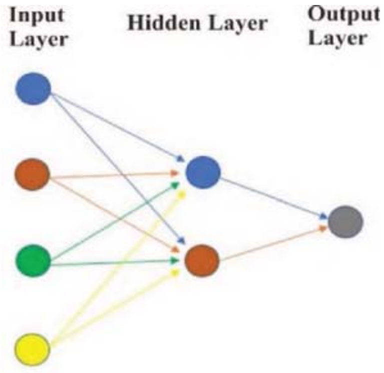

Fig. 4. BPNN Architecture 2 Neurons Hidden Layer

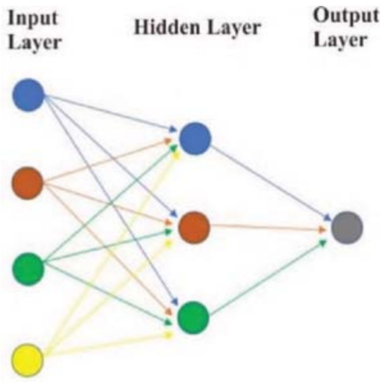

Fig. 5. BPNN Architecture 3 Neurons Hidden Layer

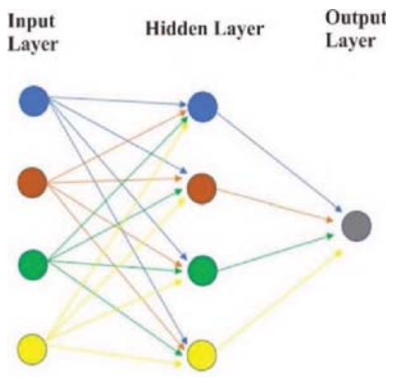

Fig. 6. BPNN Architecture 4 Neurons Hidden Layer

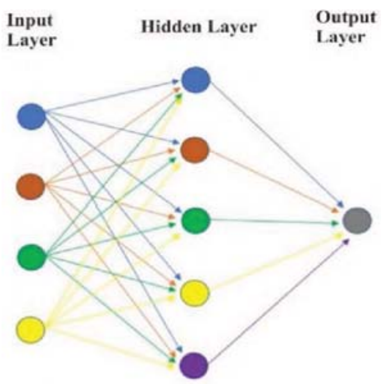

Fig. 7. BPNN Architecture 5 Neurons Hidden Layer

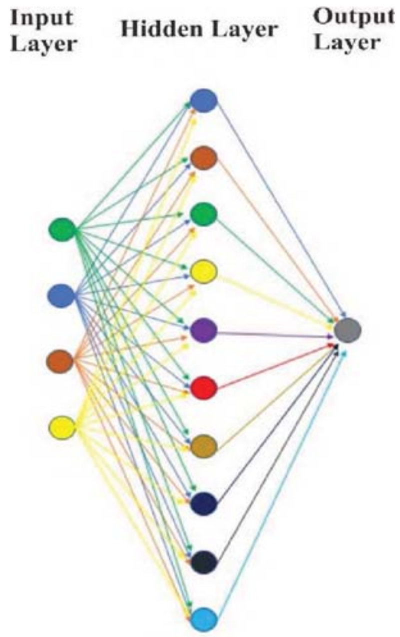

Fig. 8. BPNN Architecture 10 Neurons Hidden Layer

The hidden layers are applied to the BPNN architecture that is one hidden layer with 2, 3, 4, 5, and 10 neurons in the hidden layer. The activation function is used to connect the input layer with the hidden layer, namely the *tansig* and the hidden layer to the output layer using the *tansig*, *logsig*, and *purelin* activation functions, then look for the best activation function.

E. Single Moving Average Method

This forecasting method takes a group of observed values and then looks for the average, then using it as a prediction for the next period [18]. The term moving average is used because every time new observational data is available, the recent average is calculated and used as a prediction [19], [20]. The SMA method is necessary to calculate the average price of chili per day from 2015 to 2018 with Moving Average 2, 4, 6, and 8. Each MSE will be obtained from the moving average, and the lowest one will be selected as the best SMA forecast. SMA flowchart is presented in Fig. 9.

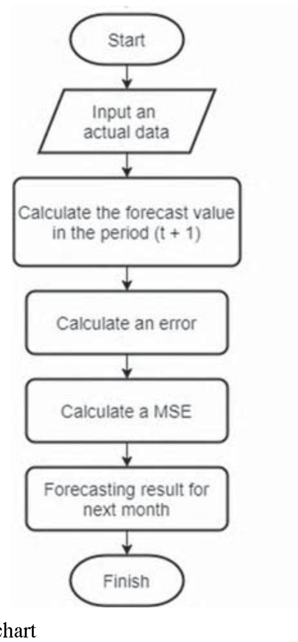

Fig. 9. SMA Flowchart

F. Calculation of Forecast Error

The accuracy of forecasting the future is crucial. If y_t is the real data for period t and Ft is the forecast for the same period, then the error can be written (4).

$$e_t = y_t - F_t \quad (4)$$

MSE is an alternative method for evaluating forecasting techniques. Each error (the difference between actual and forecast data) is squared, then added and divided by the total data (5).

$$MSE = \frac{\sum_{i=1}^{n} e_i^2}{n} \quad (5)$$

III. RESULT AND DISCUSSION

A. Results of the BPNN Method

Data of chili price is a variable to be tested using parameters of the number of hidden layers, learning rate, and activation function with a maximum epoch used of 1000 and an error tolerance of 0.01. Each of these parameters is various, according to Table II. From 1215 times of testing, the best architectural model and parameters are obtained, which is indicated by the small error rate or the MSE value. The test results with the best MSE value can be seen in Table III. Based on Table III, it can be seen that the test with the smallest MSE is always found in the *trainlm* learning function and the *purelin* activation function. The test to be performed only uses the architecture and parameters with the smallest MSE test value. Fig. 10 is presented the results and plot graphs of training and model testing.

The plot graph is using a training testing percentage of 80:20 and a learning rate of 0.9. The *trainlm* learning function and the *purelin* activation function obtained a training MSE value of 0.00023098 and MSE test of 0.00000212. It can be seen from the graph that the training results are very close to the expected results.

TABLE II. BPNN VARIABLE TESTING PARAMETERS

Parameter	
Training Testing Parameter	70:30 ; 80:20 ; 90:10
Neuron Hidden Layer	2, 3, 4, 5, 10
Learning Function	trainlm, traingd, traingdx
Learning Rate	0.1;0.2;0.3;0,4;0.5;0,6;0.7;0,8;0.9

TABLE III. BEST BPNN ARCHITECTURE MODEL AND PARAMETERS

Training Testing Percentage	Architecture Model	Learning Function (LF)	Activation Function (AF)	Learnig Rate (LR)	Mean Square Error (MSE)
70 : 30	4 – 5 – 1	trainlm	purelin	0.9	0.00133744
80 : 20	4 – 5 – 1	trainlm	purelin	0.9	0.00000212
90 : 10	4 – 5 – 1	trainlm	purelin	0.9	0.00097283

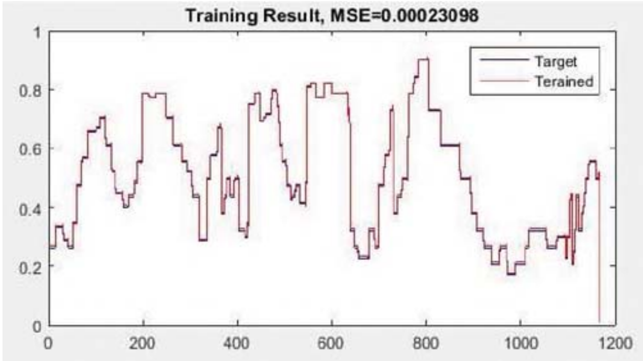

Fig. 10. Plot Graph of the BPNN Network Training 80:20

Based on the result, the best architectural model and parameters for forecasting the chili price is architecture with four input neurons, five hidden neurons, and one output neuron. Also, a learning rate of 0.9 with a maximum epoch used of 1000 with the *trainlm* learning function and *purelin* activation function. The forecast results of chili price in the 5th year or 2019 obtained an MSE value of 0.00000212, where the epoch stops in the first iteration with an execution time of 0 seconds. The final comparison of test results and actual BPNN data with the best parameters consists of 80% training data and 20% testing.

B. Results of the SMA Method

Calculating the forecast by averaging 2, 4, 6, 8, and 10 of the actual data in the previous period, the models used are Moving Average 2, 4, 6, 8, and 10 the parameters in Table IV. The results of tested parameters show that the best parameters are indicated by the small error rate or the MSE value. The test results with the best MSE value can be seen in Table V. Based on the result, it can be seen that the test with the smallest MSE is found in the MA(2). Furthermore, Fig. 11 presents the best single moving average parameter.

Based on the test results using the best parameter, it can be seen that real data and MA forecasting data are not much different. The chart presents that the expected result is close to the actual target. Based on the trial results using the best parameter (MA(2)), the MSE value is 13,142,550 or normalized to 152.2803975. This MSE value is taken from the average forecasting error in MA(2). In addition, the forecast error is obtained from each of the two previous real data averages which is calculated following Moving Average 2

TABLE IV. SMA CALCULATION PARAMETERS

Parameter	
Average	2, 4, 6, 8, 10
Moving Average	MA(2), MA(4), MA(6), MA(8), MA(10)

TABLE V. BEST SMA PARAMETERS

Moving Average	MA (2)	MA (4)	MA (6)	MA (8)	MA (10)
SSE	19188125000	26881468750	34538902778	41954304688	49225805000
MSE	13151559.29	18449875.6	23738077.51	28874263.38	33925434.18

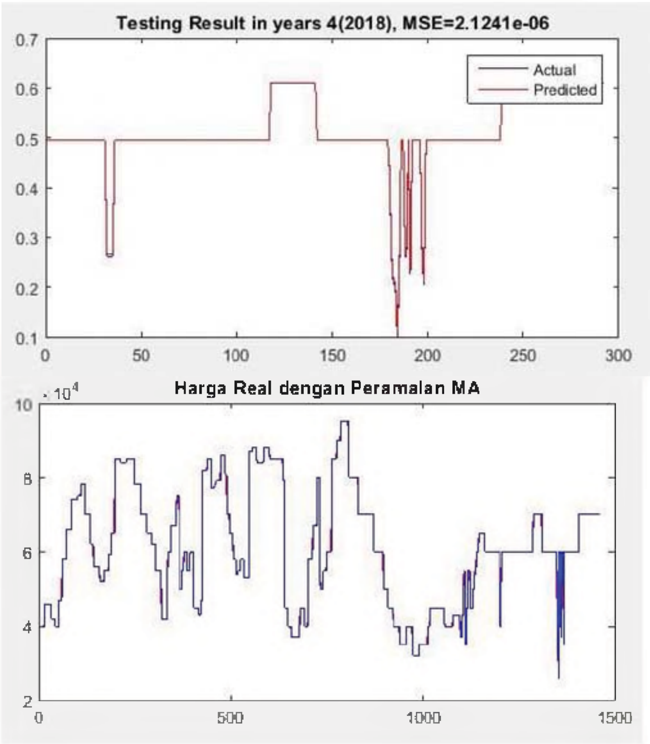

Fig. 11. Comparison Chart of Actual and Forecasting Prices

TABLE VI. MSE COMPARISON BETWEEN BPNN AND MSA

Method	MSE
BPNN	0.00000212
SMA	152.2803975

Fig. 12. Comparison Graph of BPNN and SMA Forecasting Results

C. Comparison of Prediction Results

After testing with the BPNN and SMA methods, the two methods' error rates will be compared. The comparison results of errors based on MSE can be seen in Table VI. From Table VI, it shows that the BPNN method has a much lower MSE score than SMA. The smallest MSE value was obtained in the BPNN method (0.00000212), while in SMA the normalized data was 152.2803975. To see the graph of BPNN and SMA forecasting results is presented in Fig. 12.

Based on the results, the BPNN method has a better level of accuracy than SMA. It is because the data used is a long-term and fluctuating time-series type. The SMA method is not suitable for long-term and volatile forecasting. Also, SMA is unable to cope well with trends or seasonality. It will have a better forecasting accuracy if used on constant or stable data, and the data used is more complicated. In this case study, it would be better to use the BPNN method than the SMA method.

CONCLUSION

Based on the research result, it can be concluded that forecasting results using BPNN with comparison 80%:20% of training and test data will produce the smallest MSE. It is indicated that the BPNN method has a better level of accuracy than the SMA method. Afterward, the best BPNN architectural model and parameters show that the test with the smallest MSE is always found in the *trainlm* learning function and the *purelin* activation function. By using the *trainlm* learning function, the learning rate is 0.9 with four input and five neurons in the hidden layer and one output with a 4-5-1 architecture. The maximum use of epoch is 1000, and epoch stops at the first iteration with 0 seconds of execution time. Whereas in SMA using parameters Moving Average 2, 4, 6, 8, and 10, the smallest MSE obtained is MA(2).

Furthermore, the SMA method produces a greater MSE than the BPNN method. It means that the SMA method is not suitable for long-term and volatile forecasting. Besides, SMA is also unable to cope well with trends or seasonality. Not all data can be used as approximate forecast data. Some data require certain variables or parameters to fulfill the method used. Further research can be redeveloped by a suitable method with fluctuating data following the chili price characteristics and adding other variables. This research can also be developed using other neural network learning methods and can be compared with different statistical methods.

REFERENCES

[1] M. F. Syuaib, "Sustainable agriculture in Indonesia: Facts and challenges to keep growing in harmony with environment," *Agric. Eng. Int. CIGR J.*, vol. 18, no. 2, pp. 170–184, 2016.

[2] J. Mariyono, "Moving to commercial production: a case of intensive chili farming in Indonesia," *Dev. Pract.*, vol. 27, no. 8, pp. 1103–1113, 2017.

[3] I. Y. G. Ustriyana and I. A. L. Dewi, "Analysis of perception of chili farmers on sustainable development," *Am. J. Sustain. Agric.*, vol. 11, no. 4, pp. 23–29, 2017.

[4] M. Sativa, H. Harianto, and A. Suryana, "Impact of red chilli reference price policy in Indonesia," *Int. J. Agric. Syst.*, vol. 5, no. 2, pp. 120–139, 2017.

[5] S. Hasmita, F. Nhita, D. Saepudin, and A. Aditsania, "Chili Commodity Price Forecasting in Bandung Regency using the Adaptive Synthetic Sampling (ADASYN) and K-Nearest Neighbor (KNN) Algorithms," in *2019 International Conference on Information and Communications Technology (ICOIACT)*, 2019, pp. 434–438.

[6] G. Hegde, V. R. Hulipalled, and J. B. Simha, "A Study On Agriculture Commodities Price Prediction and Forecasting," in *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)*, 2020, pp. 316–321.

[7] S. Makridakis, E. Spiliotis, and V. Assimakopoulos, "Statistical and Machine Learning forecasting methods: Concerns and ways forward," *PLoS One*, vol. 13, no. 3, p. e0194889, 2018.

[8] S. J. Taylor and B. Letham, "Forecasting at scale," *Am. Stat.*, vol. 72, no. 1, pp. 37–45, 2018.

[9] J. Kleinberg, H. Lakkaraju, J. Leskovec, J. Ludwig, and S. Mullainathan, "Human decisions and machine predictions," *Q. J. Econ.*, vol. 133, no. 1, pp. 237–293, 2018.

[10] H. Haviluddin and I. Tahyudin, "Time series prediction using radial basis function neural network," *Int. J. Electr. Comput. Eng.*, vol. 5, no. 4, pp. 765–771, 2015.

[11] R. Alfred, "A genetic-based backpropagation neural network for forecasting in time-series data," in *2015 International Conference on Science in Information Technology (ICSITech)*, 2015, pp. 158–163.

[12] K. B. Debnath and M. Mourshed, "Forecasting methods in energy planning models," *Renew. Sustain. Energy Rev.*, vol. 88, pp. 297–325, 2018.

[13] R. R. Asaad and R. I. Ali, "Back Propagation Neural Network (BPNN) and sigmoid activation function in multi-layer networks," *Acad. J. Nawroz Univ.*, vol. 8, no. 4, pp. 216–221, 2019.

[14] B. Priambodo, W. F. Mahmudy, and M. A. Rahman, "Earthquake Magnitude and Grid-Based Location Prediction using

Backpropagation Neural Network," *Knowl. Eng. Data Sci.*, vol. 3, no. 1, pp. 28–39, 2020.

[15] M. M. Ali, M. Z. Babai, J. E. Boylan, and A. A. Syntetos, "Supply chain forecasting when information is not shared," *Eur. J. Oper. Res.*, vol. 260, no. 3, pp. 984–994, 2017.

[16] H. Aini, H. Haviluddin, E. Budiman, M. Wati, and N. Puspitasari, "Prediksi Produksi Minyak Kelapa Sawit Menggunakan Metode Backpropagation Neural Network," *Sains, Apl. Komputasi dan Teknol. Inf.*, vol. 1, no. 1, pp. 24–33, 2019.

[17] K. Wong, A. P. Wibawa, H. S. Pakpahan, A. Prafanto, and H. J. Setyadi, "Prediksi tingkat inflasi dengan menggunakan metode backpropagation neural network," *Sains, Apl. Komputasi dan Teknol. Inf.*, vol. 1, no. 2, pp. 8–13, 2019.

[18] M. B. C. Khoo and P. W. Yap, "Joint monitoring of process mean and variability with a single moving average control chart," *Qual. Eng.*, vol. 17, no. 1, pp. 51–65, 2004.

[19] S. Siami-Namini, N. Tavakoli, and A. S. Namin, "A comparison of ARIMA and LSTM in forecasting time series," in *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA)*, 2018, pp. 1394–1401.

[20] A. Tealab, H. Hefny, and A. Badr, "Forecasting of nonlinear time series using ANN," *Futur. Comput. Informatics J.*, vol. 2, no. 1, pp. 39–47, 2017.