Jurnal Sains Komputer & Informatika (J-SAKTI)
Volume 7 Nomor 2, September 2023, pp. 1015-1024
ISSN: 2548-9771/EISSN: 2549-7200
https://tunasbangsa.ac.id/ejurnal/index.php/jsakti

Application of the LSTM Algorithm in Predicting Urea Fertilizer Production at IIB Plant PT. Pupuk Sriwidjaja Palembang

Aziz Awaludin[1], Ferdiansyah[2], Andri[3], Tri Oktarina[4]
[1,2,3,4]Universitas Bina Darma, Palembang, Indonesia
e-mail: [1]azawldn@gmail.com, [2]ferdi@binadarma.ac.id, [3]andri@binadarma.ac.id, [4]tri_oktarina@binadarma.ac.id

Abstract
PT. Pupuk Sriwidjaja Palembang is a pioneer of fertilizer manufacturers in Indonesia. One of the plants at PT. Pupuk Sriwidjaja Palembang, namely the IIB urea plant, has been operating normally since 2017, thereby the data of production results has been collected for more than five years (time series data). The collected data can be used to make predictions of future production using the LSTM (Long Short Term Memory) model. LSTM is an artificial neural network architecture that is suitable for processing sequential data. The research objective to be achieved is to produce a production prediction model using LSTM modeling. Data collected over five years was divided into training data and testing data through data composition trials. The LSTM model training was carried out with a training data composition of 70% of the total data, batch size 64, and epoch 200. Then testing was carried out with data testing as much as 30% of the total data using RMSE and MAPE as model quality assessment parameters. Based on test results, the LSTM model is able to predict production with an RMSE of 11.08 and a MAPE of 6.39%.

***Keywords**: LSTM, Urea Fertilizer, Production, Predictions*

1. INTRODUCTION

PT. Pupuk Sriwidjaja Palembang (PT. Pusri) was established on December, 24 1959 in Palembang, South Sumatra as a pioneer of fertilizer manufacturers in Indonesia. PT. Pusri started their business with the main aim of implementing and supporting government policies and programs in the economic and national development sectors, especially in the fertilizer and other chemical industries. One of the main products of PT. Pusri is urea fertilizer, which is a chemical fertilizer that contains high levels of Nitrogen (N). Current advances in information technology are both a challenge and an opportunity for companies to develop their business and continue to compete by innovating in line with current technology developments. Companies have to optimize their own resources, including the data. Data is a value that represents a description of an object or event. Information is the result of data processing in a form that is more useful and more meaningful for the recipient which describes real events that are used for decision making. Data is more historical, while information has a higher level; more dynamic, and has very important value [1].

One of the urea fertilizer plant at PT. Pusri, namely the IIB plant, was founded in 2015, started producing urea fertilizer since 2017, and operated normally since 2018. As a fertilizer producing company, production activities at PT. Pusri takes place every day for 24 hours without any rest hour. Urea fertilizer production results are documented every day, so the data is considered as time series data. Time series data refers to the results of a certain observation process, taking certain samples in a period with the same time interval. The essence of time series analysis is to discover laws from data and predict future values based on historical observations, which can provide a reference and basis for decision making [2]. The use of time series data that has been recorded in large quantities can still be optimized to design prediction models that can be used to determine targets and evaluate factory performance in producing urea fertilizer. Prediction or forecasting is an attempt to estimate something that will happen in the future by utilizing various relevant information at previous times through a scientific method. [3].

Prediction models can be created using deep learning. Deep Learning is a branch of machine learning that uses Artificial Neural Networks (ANN) as its basis. Deep Learning is a network consisting of several layers which are a collection of nodes where a calculation process occurs. An input node will be calculated with the weight of the data, then the results of the calculation will go through a stage called the node activation function, to determine how far the signal goes further through the network. [4].

One model that can be used to process time series data, namely Long Short Term Memory networks (LSTM) is an evolution of the Recurrent Neural Network (RNN) architecture, which was first introduced by Hochreiter & Schmidhuber (1997). RNN is not able to process data learning by connecting long information, because the old stored memory will become increasingly useless because the data is overwritten or replaced by new memory, this problem was discovered by Bengio, et al [5]. The LSTM model has advantages in the learning process, because LSTM is able to remember longer past time series data than RNN [6]. Based on the background above, the production prediction model developed in this research is a model based on the Long Short-Term Memory algorithm, because the data used in this research is time series data.

2. RESEARCH METHODOLOGY

The method used in this research is the descriptive method, which systematically describing facts and information based on historical data by applying the LSTM (Long Short-Term Memory) algorithm. Data collection was carried out by making observations at the research site. The literature study was carried out by collecting information from scientific works, books and journals to be used as a reference for this research.

2.1. Long Short-Term Memory Neural Networks

LSTM neural networks are well suited for classifying, processing, and making predictions based on time series data because there may be a dearth of unknown durations between important events in a time series [7].

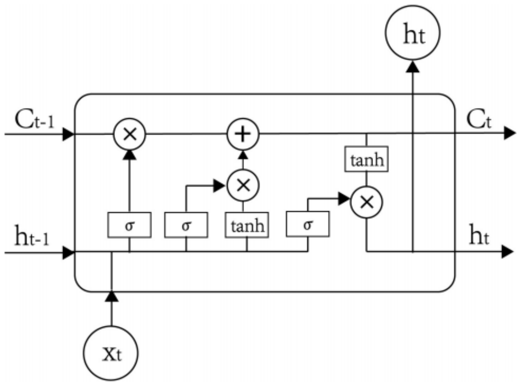

Figure 1. LSTM Architecture

LSTM has 3 gates, namely input gate, forget gate and output gate. Based on research [8], the computing process in LSTM is carried out in the following stages. The value of an input can only be stored in cell state if permitted by the input gate. The value at the input gate, i_t and the candidate value of the cell state, $\bar{C}_t$, at the time step, t, are calculated using the following equation:

$$i_t = \sigma(W_i x_i + U_i h_{t-1} + b_i), \quad (1)$$
$$\bar{C}_t = tanh(W_c x_t + U_c h_{t-1} + b_c), \quad (2)$$

where W, U, b represent matrix weights and bias respectively. The weight of the unit state is set by the forget gate, f_t, the forget gate value is calculated as follows:

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

Through this process, the new cell state, C_t, is calculated as follows:

$$C_t = i_t \bar{C}_t + f_t\, C_{t-1}. \quad (4)$$

With the new cell state, the value of the gate output, o_t, is calculated as follows:

$$o_t = \sigma(W_o x_i + U_o h_{t-1} + V_o C_t + b_o). \quad (5)$$

The last output value of a cell is defined as the hidden state, h_t, calculated as follows:

$$h_t = o_t tanh(C_t). \quad (7)$$

2.2. Related Researches

Based on research [9], regarding short-term load forecasting of the Batu city electricity system using Long Short Term Memory, researchers carried out 5 simulations of the proportion of training data and testing data, namely 50:50, 60:40, 70:30, 80:20, and 90:10 using two data source; Penyulang Panorama and Penyulang Wastra Indah. Researchers used the Autoregressive Integrated Moving Average (ARIMA) and LSTM models as comparisons. It is known that from two different data sources, the best predictions for both are produced by the largest proportion of training data, namely 90:10.

In research conducted by [10], researchers used sales data from PT. Metiska Farma between 2017 and 2019 to create a prediction model using LSTM. Researchers conducted trials on 3 dataset compositions, namely 70:30, 80:20, and 90:10 as well as 2 trials of interval values, namely [-1,1] and [0,1] to find the combination that can produce the best predictions. Good. The results of research conducted by researchers show that the best method for using LSTM is with a composition of 90% training data and 10% test data, interval range [-1,1] and 1500 epochs.

In research conducted by [11], researchers predicted Bitcoin from the Yahoo Finance stock market with LSTM. Researchers use varying numbers of epochs in training training data to find out the optimal number of epochs that should be used when training training data to obtain the best prediction results. Tests were carried out on 8 variations of epoch values, namely; 10, 100, 1000, 200, 400, 800, 2000, 5000. The results obtained, namely the use of epoch 500, produce a model with the lowest RMSE value.

In research conducted by [12], researchers used the gate recurrent unit (GRU) and bidirectional–LSTM (Bi-LSTM) hybrid model to predict cryptocurrency prices to improve the accuracy and normalize the root mean square error (RMSE). Researchers tested the model with epochs 10, 250, and 400 to find out the model with the smallest RMSE value for the prediction of four cryptocurrencies (Bitcoin, Ethereum, Ripple, and Binance).

2.3. Research Stages

The research stages carried out are as follows

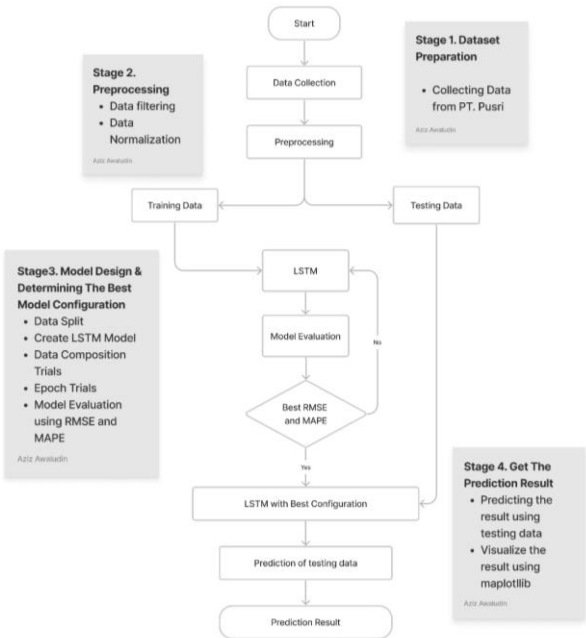

Figure 2. Research Stages

1) Stage 1: Dataset Preparation

The data used as a dataset is urea production results data at the IIB factory, PT. Pupuk Sriwidjaja Palembang which considered as time series data with over a period of five years starting from September, 29 2018 – July, 31 2023 with 1707 rows and 11 columns of data.

TANGGAL	UJI KINERJA		PRODUKSI UREA			PRODUKSI UREA (TON)	KONSUMSI CO2 (TON)	KONSUMSI NH3 (TON)	RATE PRODUKSI (%)	RATE OPERASI (%)	WIS 301
	RASIO CO2 (TON/TON)	RASIO NH3 (TON/TON)	BASE CO2 (TON)	BASE NH3 (TON)	BASE WIS-301 (TON)						
29 Sep 2018	0,77	0,57	3.002,33	3.021,45	3.071,88	2.794,54	2.296,78	1.658,94	101,62	115,57	3.071,88
30 Sep 2018	0,77	0,57	3.008,32	3.016,09	3.037,50	2.789,25	2.301,36	1.655,52	101,43	115,97	3.037,50
01 Okt 2018	0,77	0,57	3.004,19	3.019,05	3.026,50	2.812,64	2.298,20	1.669,07	102,28	115,91	3.026,50
02 Okt 2018	0,77	0,57	3.002,83	3.017,95	3.023,88	2.811,55	2.297,16	1.668,02	102,24	115,82	3.023,88
03 Okt 2018	0,77	0,57	3.008,71	3.015,54	3.003,88	2.809,15	2.301,67	1.666,97	102,15	115,93	3.003,88
04 Okt 2018	0,77	0,57	3.006,43	3.014,67	3.035,88	2.808,29	2.299,92	1.665,51	102,12	115,71	3.035,88
05 Okt 2018	0,77	0,57	2.996,13	3.008,65	3.017,25	2.802,30	2.292,04	1.662,41	101,90	115,67	3.017,25
06 Okt 2018	0,77	0,57	3.013,97	3.028,02	3.014,00	2.821,56	2.305,69	1.674,18	102,60	116,27	3.014,00
07 Okt 2018	0,77	0,57	3.017,21	3.031,74	3.055,38	2.825,26	2.308,17	1.675,67	102,74	116,28	3.055,38
08 Okt 2018	0,77	0,57	3.022,85	3.034,26	3.040,00	2.827,77	2.312,48	1.677,82	102,83	116,35	3.040,00
09 Okt 2018	0,77	0,57	3.015,64	3.024,30	3.049,63	2.817,86	2.306,97	1.672,09	102,47	116,21	3.049,63
10 Okt 2018	0,77	0,57	3.012,60	3.020,25	3.027,63	2.813,84	2.304,64	1.670,39	102,32	115,95	3.027,63
11 Okt 2018	0,77	0,57	3.010,20	3.025,72	3.047,75	2.819,28	2.302,80	1.674,55	102,52	116,06	3.047,75
12 Okt 2018	0,77	0,57	3.016,53	3.027,25	2.993,13	2.820,80	2.307,64	1.673,52	102,57	116,22	2.993,13
13 Okt 2018	0,77	0,57	1.884,54	1.538,75	1.537,44	1.340,93	1.441,67	795,76	48,76	63,14	1.537,44
14 Okt 2018	0,77	0,57	2.778,61	2.705,67	2.560,00	2.501,08	2.125,64	1.484,25	90,95	103,50	2.560,00
15 Okt 2018	0,77	0,57	2.962,88	2.971,21	2.904,63	2.765,08	2.266,61	1.640,91	100,55	114,51	2.904,63
16 Okt 2018	0,77	0,57	3.000,12	3.005,58	2.857,50	2.799,25	2.295,09	1.661,20	101,79	116,02	2.857,50

Figure 3. Raw Data

2) Stage 2: Preprocessing

At the preprocessing stage, unnecessary data is eliminated using the filter function.

TANGGAL	PRODUKSI UREA (TON)
2018-09-29	2794.54
2018-09-30	2789.25
2018-10-01	2812.64
2018-10-02	2811.55
2018-10-03	2809.15
...	...
2023-07-27	2884.39
2023-07-28	2889.95
2023-07-29	2884.70
2023-07-30	2897.80
2023-07-31	2893.08

1707 rows × 1 columns

Figure 4. Dataset

The data is converted into array and normalized using MinMaxScaler.

3) Stage 3: Model Design and Determining The Best Model Configuration

The model designed is the Long Short Term Memory (LSTM) model which was developed using Python and Keras as the framework. The model consists of 3 LSTM layers and 2 Dense layer; 256 units in the first LSTM layer, 128 units in the second LSTM layer, 64 units in the third LSTM layer, 32 Units in the first dense layer, and 1 unit in the last dense layer. The model is compiled using the compiler and Adam optimizer.

Data that has gone through the preprocessing stage is divided into two parts, namely training data and testing data. Model evaluation was carried out using two assessments, using Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). RMSE is one of the assessment parameters used to evaluate how well a model can predict continuous values. A low RMSE value indicates that the variation in values produced by a prediction model is close to the variation in values in the actual data. RMSE calculates how different a set of values are. The smaller the RMSE value, the better the model is at making predictions [13].

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(x-\hat{x})^2} \tag{8}$$

RMSE = root mean squared error value
N = Number of observations
xi = Actual value
$\hat{x}$i = Predicted value

Mean Absolute Percentage Error (MAPE) is a prediction model assessment parameter that describes the percentage error in predictions of actual results during a certain period. MAPE is calculated using the absolute error in each period divided by the actual observed value [14] [15].

$$MAPE = \frac{100}{N}\sum\left|\frac{x-\hat{x}}{x}\right| \tag{9}$$

The MAPE value can be interpreted or interpreted into 4 categories [9], that is:
<10% = Very accurate
10-20% = Good
20-50% = Reasonable
> 50% = Inaccurate
The smaller the MAPE value, the smaller the error in the estimation results.
The trials were carried out on training data and testing data to find out the best data composition for making predictions, and trials of varying epoch values were carried out using the data composition that had been determined in the data composition trial. Based on previous research regarding determining data composition and testing various epoch values, the determination of the composition of training data and testing data in this research is determined through trials on 3 compositions of training data and testing data, namely 70:30, 80:20, and 90: 10. Meanwhile, for testing various epoch values in this research, trials were carried out on epoch values 10, 100, 200, 500, and 1000.

The composition of training data and testing data with the lowest RMSE and MAPE values will be used in trials with varying epoch values. The model that has the best RMSE and MAPE values in the epoch trial will be used as the prediction model for this research.

4) Stage 4 : Get The Prediction Result

At this stage the model with the best configuration is used to make predictions. Prediction results will be compared with actual data by visualizing it using maplotlib.

3. RESULTS AND DISCUSSION

3.1. Model Evaluation

Model evaluation is carried out to determine the model with the best prediction results based on the evaluation parameter values used, namely RMSE and MAPE. Two trial processes were carried out in determining the training model configuration in this research, namely the data composition trials , and trials on varying epoch values.

In the testing process, training data and testing data were carried out with 3 compositions, namely 70:30, 80:20, and 90:10 using batch size 64 and epoch 200 for each composition. The trial results were evaluated using RMSE and MAPE. The following results were obtained;

Table 1. Data Composition Trial Results

No	Data Composition	RMSE	MAPE
1	70:30	11.08	6.39%
2	80:20	32.89	6.90%
3	90:10	31.51	8.35%

Based on table 1, data composition of 70:30 has the lowest RMSE and MAPE values, so a training data composition of 70% and testing data of 30% is used as the data composition in this research. Next, various epoch value trials. The training was carried out on data with batch size 64 and varying epochs; 10, 100, 200, 500, 1000. The trial results were evaluated using RSME and MAPE with the following results;

Table 2. Various Epoch Value Trial Results

No	Epoch	RMSE	MAPE
1	10	83.30	8.57%
2	100	43.00	7.68%
3	200	11.08	6.39%
4	500	86.68	6.73%
5	1000	99.07	7.31%

Based on the various epoch value trial results in table 2, the model that went through the training stage with epoch 200 produced the lowest RMSE and MAPE values. So the model in this research used a data composition with 70% training data, 30% testing data, batch size 64, and epoch 200.

3.2. Prediction Results

Urea production results predicted using the LSTM model which have been evaluated using RMSE and MAPE in the trials above can be seen in Figure 5 and Figure 6.

TANGGAL	PRODUKSI UREA (TON)	predictions
2022-02-21	2852.205	2882.901367
2022-02-22	2939.317	2812.376465
2022-02-23	2960.065	2900.951416
2022-02-24	2952.161	2874.843262
2022-02-25	2798.526	2876.718262
...	...	...
2023-07-27	2884.390	2832.923340
2023-07-28	2889.950	2825.062988
2023-07-29	2884.700	2835.323242
2023-07-30	2897.800	2831.641113
2023-07-31	2893.080	2846.667725

512 rows × 2 columns

Figure 5. Prediction Results Compared to Urea Production Data

Figure 5 shows the prediction results compared to urea production data. The urea production data were the testing data containing 30% of the total data after it gone through the data split process in the second research stages. The testing data were the data from February, 21 2022 to July, 31 2023. From the comparison in figure 5, it is known that the model is able to do the predictions of the urea productions.

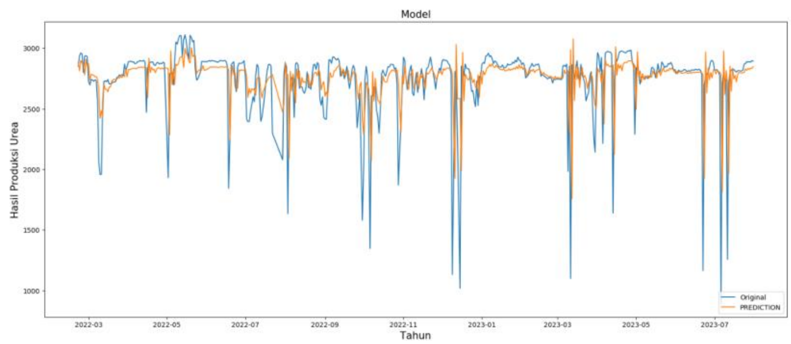

Figure 6. Comparison Graph : Prediction vs Urea Production Data

Figure 6 shows the comparison graph between prediction results and urea production data. From the comparison graph in figure 6, it is known that the model is not able to do a good prediction where the data suddenly gone too low from its latest trends. Several factors that can cause the urea production to be very low, namely the possibility of the plant maintenance, plant cleaning, or the occurrence of unexpected conditions such as the lost of power supplies caused by some incidents at the internal power plants.

4. CONCLUSION

Determining the composition of training data and testing data, as well as choosing the right epoch when training the model, greatly influences the quality of the prediction results produced by the model. Based on the results of testing data composition trials and various epoch value trials, it is known that the best configuration for the model in this research is using training data as much as 70% of the total data, testing data as much as 10% of the total data, batch size 64, and epoch 200. Based on The results of the model evaluation using RMSE and MAPE, obtained an RMSE value of 11.08 and a MAPE of 6.39%, so it can be concluded that the model is able to do the predictions. However, based on the comparison graph of predicted results with actual data in Figure 6, It is known that the model cannot predict the possibility of the plant experiencing a sudden decline in production.

Better prediction results can be obtained if the plant conditions are always stable. This model can also be used to produce predictions that can be used as a reference for PT. Pusri to determine production targets in accordance with recent plant performance.

REFERENCES

[1] C. A. Pamungkas, "BAB I Pendahuluan," in *Pengantar dan Implementasi Basis Data*, Yogyakarta, DEEPUBLISH, 2017, p. 1.

[2] Z. Liu, Z. Zhu, J. Gao and C. Xu, "Forecast Methods for Time Series Data: A Survey," *IEEE Access,* vol. 9, pp. 91896-91912, 2021.

[3] E. Hartato, D. Sitorus and a. A. Wanto, "Analisis Jaringan Saraf Tiruan Untuk Prediksi Luas Panen Biofarmaka Di Indonesia," *semanTIK,* vol. 4, no. 1, pp. 49-56, 2018.

[4] M. Rizki, S. Basuki and Y. Azhar, "Implementasi Deep Learning Menggunakan Arsitektur Long Short Term Memory Untuk Prediksi Curah Hujan Kota Malang," *REPOSITOR,* vol. 2, no. 3, pp. 331-338, 2020.

[5] A. Arfan and Lussiana, "Prediksi Harga Saham Di Indonesia Menggunakan Algoritma Long Short -Term Memory," *Seminar Nasional Teknologi Informasi dan Komunikasi STI&K (SeNTIK),* vol. 3, no. 1, pp. 225-230, 2019.

[6] M. Mukhlis, A. Kustiyo and A. Suharso, "Peramalan Produksi Pertanian Menggunakan Model Long Short-Term Memory," *BINA INSANI ICT JOURNAL,* vol. 8, no. 1, pp. 22-32, 2021.

[7] A. Khumaidi, R. Raafi'udin and I. P. Solihin, "Pengujian Algoritma Long Short

Term Memory untuk Prediksi Kualitas Udara dan Suhu Kota Bandung," *Jurnal Telematika,* vol. 15, no. 1, pp. 13-18, 2020.

[8] H. Chung and K.-s. Shin, "Genetic Algorithm-Optimized Long Short-Term Memory Network for Stock Market Prediction," *Sustainability,* vol. 10, no. 3765, pp. 1-18, 2018.

[9] H. Purnomo, H. Suyono and R. N. Hasanah, "Peramalan Beban Jangka Pendek Sistem Kelistrikan Kota Batu Menggunakan Deep Learning Long Short-Term Memory," *Transmisi : Jurnal Ilmiah Teknik Elektro, 23, (3),* vol. 23, no. 3, pp. 97-102, 2021.

[10] L. Wiranda and M. Sadikin, "Penerapan Long Short Term Memory Pada Data Timeseries Untuk Memprediksi Penjualan Produk PT. Metiska Farma," *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI,* vol. 8, no. 3, pp. 184-196, 2019.

[11] Ferdiansyah, S. H. Othman, R. Z. M. Radzi, D. Stiawan, Y. Sazaki and U. Ependi, "A LSTM-Method for Bitcoin Price Prediction: A Case Study Yahoo Finance Stock Market," in *International Conference on Electrical Engineering and Computer Science (ICECOS)*, Batam, 2019.

[12] Ferdiansyah, S. H. Othman, R. Z. M. Radzi, D. Stiawan and T. Sutikno, "Hybrid gated recurrent unit bidirectional-long short-term memory model to improve cryptocurrency prediction accuracy," *IAES International Journal of Artificial Intelligence (IJ-AI),* vol. 12, no. 1, pp. 251-261, 2023.

[13] R. J. Hyndman and A. B. Koehler, "Another look at measures of forecast accuracy," *International Journal of Forecasting,* vol. 22, no. 4, pp. 679-688, 2006.

[14] M. L. Ashari and M. Sadikin, "Prediksi Data Transaksi Penjualan Time Series," *Jurnal Nasional Pendidikan Teknik Informatika : JANAPATI,* vol. 9, no. 1, pp. 1-10, 2020.

[15] C. D. Lewis, Industrial and Business Forecasting Methods: A Practical Guide to Exponential Smoothing and Curve Fitting, London: Butterworth Scientific, 1982.