

Backpropagation Neural Network for Predict Sugarcane Stock Availability



Siti Nurhasanah¹, Indra Ranggadara², Ifan Prihandi³, Anita Ratnasari⁴

^{1,2,3,4}Faculty of Computer Science, MercuBuana University, Jakarta, Indonesia

41816110141@student.mercubuana.ac.id¹, indra.ranggadara@student.mercubuana.ac.id²

ifan.prihandi@student.mercubuana.ac.id³, anita.ratnasari@student.mercubuana.ac.id⁴

ABSTRACT

National sugar consumption from 2013 to 2016 experienced an upward trend of 6.14%, especially for 2016, demand increased by 5.08% from the previous year[1]. Increment consumption of this sugar is inversely proportional to production results domestic sugarcane. Domestic sugarcane yields each year has decreased due to the problem of sugarcane on-farm or off the farm. Hence the government is trying to import sugarcane raw materials to fulfill national sugar needs. In importing raw materials sugarcane, the government needs predictive data for sugarcane production and the national sugar needs that will come used as a reference for the amount of raw sugarcane that needs to be imported. This study aims to estimate the results of sugarcane production based on historical data from 2017 to 2019. After making further predictions, it is necessary to evaluate the prediction model used. Method research used to predict production yield sugarcane, namely the Backpropagation Neural Network. Furthermore, the evaluation of the model used uses the Mean Absolute Percentage Error (MAPE). MAPE delivers a hint of how large the forecast error is compared with the series's real value. The following result 97% predictable production and 3% MAPE yield, then it can be said that the result of the method Backpropagation Neural Network has modeling capabilities decent forecast.

Key words : Sugarcane, Stock, Prediction, Backpropagation Neural Network

1. INTRODUCTION

Increased population and the development food industry, which requires raw materials in the form of sugar, makes national sugar needs increase every year. Sugar is an agricultural commodity Indonesia has designated as a particular commodity in the World Trade Organization (WTO) negotiation forum and rice, corn, and soybeans. With consideration key to strengthening food security and quality of life in rural areas, Indonesia seeks to increase production in the country, including launching a target of self-sufficiency in sugar, that is until now it has not been achieved. Apart from not optimal factors, such conditions support domestic sugar production (on-farm and off-farm), and national sugar consumption is still high [2].

High national sugar consumption can be demonstrated through data from the 2017 Socio-Economic Survey that the average consumption of sugar per capita in a month is 5,212 ounces [1]. Indonesia's population projection in 2017 is amounting to 261,8609 million people, so it is estimated that granulated consumption sugar in 2017 was 3.376 million tons while the area sugarcane plantation area from 2014 to 2017 continues decreased as well as the production of sugar in Indonesia. The supply of sugar that the production cannot fulfill domestic activity raises sugar imports or raw sugar imports. In importing sugarcane raw materials, the government requires predictive data for sugarcane production, and future national sugar needs to be referred to the amount of raw sugarcane that needs to be imported. This matter did so that the amount of raw sugarcane is imported effectively and according to national needs, which are not fulfilled by farmers' sugarcane production[1]. Fixed farmer sugarcane is used as the first option to be marketed, and imported sugarcane as a backup stock—the research method used to predict sugarcane yield, namely by Backpropagation Neural Network.

Neural Network Backpropagation Algorithm is already widely applied to predict something with analyzing historical data[3]. This research algorithm Backpropagation Neural Network is used for historical training data from 2017 to 2019 and is expected to generate predicted values for 2020. The next evaluation of the model used is using Mean Absolute Percentage Error (MAPE)[4]. MAPE provides a clue show significant the forecast error is compared to the value actually from the series. The current MAPE value will be obtained the percentage level of accuracy from the prediction of production cane.

The contribution of this research is to make a system that can predict sugarcane yield from data sugarcane production in previous years, not only rely on manual calculations only. By knowing the prediction of sugarcane production, the prediction of sugarcane production is expected to endeavor so that the production of sugarcane is higher than the predictions are there and can become self-sufficient in sugar nations.

2. STUDY LITERATURE AND PREVIOUS RESEARCH

2.1 Sugarcane

Sugarcane plant (*Saccharum officinarum* L.) plants are of high economic value because of the primary raw material in sugar manufacture. The sugarcane plant contains sap, which can be processed into crystal sugar. Sugar is a simple carbohydrate that becomes the primary source of energy and trading commodities—the most widely traded in the form of solid sucrose crystals. Sugar is used to change the taste to sweet and the state of food or drink. Simple sugar, like glucose (produced from sucrose by enzymes or acid hydrolysis), storing energy used by the cell[5].

2.2 Stock

In carrying out its production activities, each company, either service company or company manufacturer, must hold inventory. A company that is not having inventory will be exposed to risk two risk, namely a shortage of products at a time to make customer demand is not met, but supplies are excess will make storage costs relatively large. Therefore, inventory must be well managed because it affects production and sales activities. According to Assauri (1980), the definition of inventory is assets that include the company's property with the intent to be sold within a business period regularly, or in-stock items quality/production process, or raw material inventory waiting for its use in a production process. Meanwhile, according to Rangkuti (2004), inventory is an asset that includes the company's property to be sold within a specific business period or inventory of goods that are still in progressor the production process, or inventory of raw materials waiting for its use in a production process[6]. It can be concluded from the above definitions that the company has inventory because inventory is a costly asset; assets in this company can be immediately resold or further processed at a time-specific period. Supplies are essential for a company because the inventory connects one operation to the next operation in sequence, making it delivered to the consumer. Inventory can be optimized by holding better production planning and management of optimal supply[7].

2.3 PreviousResearch

Previous research is a reference in research making this time. Neural Backpropagation Network has been widely applied to predict various in the field of information systems, one of which is in the agricultural sector. Implementation of the backpropagation algorithm in information systems can predict land use [8]. Algorithm implementation backpropagation in information systems can predict the wind [9]. The Backpropagation algorithm has applied to predict rice production; thus, it can be anticipated in advance so that rice production can be obtained become self-sufficient in food [10]. The backpropagation Algorithm can produce information systems that can predict the oil palm commodity price as an essential reference in taking company policy [11]. Algorithm Backpropagation has been applied to predict rice productivity

outcomes to optimize agricultural techniques [12]. The Backpropagation algorithm has been implemented to predict the yield of palm oil to facilitate future production planning [13]. The Backpropagation algorithm has been implemented to predict the yield of rubber commodity crops [14]. Besides matters relating to agriculture, the backpropagation algorithm can be used for the socio-cultural field, backpropagation also can identify a character[15][16].

Based on literature studies from several previously Neural Network Backpropagation Algorithm studies, it is considered good enough to predict sugarcane production. For this reason, this research intends to contribute by creating a system that can predict outcomes sugarcane from previous years' production data, not only rely on manual calculations only. By knowing the prediction of sugarcane production is expected to endeavor at the beginning so that the production of sugarcane is higher than the yield predictions are there and can become self-sufficient in sugar national.

3. RESEARCH METHOD

This section describes the method in detail Backpropagation Neural Network, measurement of simulation results predictions, and data used in this study.

3.1 Data Collection

This research took place at PT. Perkebunan Nusantara X, particularly data on people's sugarcane (sugarcane from partner farmers of PTPN X).

The data taken is 2017-2019. Every year, data is only taken from May to October because the sugarcane milling process is only carried out in that month. There are 18 data with three categories, namely data on the area of milled garden land, the number of milled sugarcane, and the yield of milled sugarcane.

3.2 Backpropagation neural network

Steps that need to be done toBPNN training is [17]:

Table 1:Data Collection

Month	Field area milled (Ha)		Amount of Sugar Cane Milled (Ton)		Yield (%)		Amount of Sugarcane Milled (Ton)
	2017	2018	2017	2018	2017	2018	2019
May	825,37	3.881,22	67.586,50	299.790,80	7,63	7,86	383.762,10
June	3.791,14	7.256,45	391.953,42	545.004,20	7,61	7,91	611.776,60
July	13.441,66	14.257,00	992.498,60	1.090.068,60	7,71	7,96	1.079.757,70
August	13.635,45	13.136,66	1.130.260,10	992.634,00	8,05	8,27	860.588,80
September	13.653,14	13.035,81	980.931,30	780.295,80	8,29	8,46	517.666,70
October	8.830,32	1.347,01	536.853,68	99.484,10	7,95	8,00	16.514,50

1. Initialize the weights with small random numbers.
2. As long as the stop conditions are not met, work onsteps 3 - 8.

Stage 1: Feed Forward (Feed forward)

3. Each input unit($x_i, i = 1, \dots, n$) revceiced input clues x_i and pass on the units hidden(*hiddenlayer*).

4. Each hidden unit ($z_j, j = 1, \dots, p$) adds up the weight of the input signal with equation (1).

$$Z_{in_jk} = v_{0j} + \sum_{i=1}^n x_i v_{ij} \quad (1)$$

Which, Z_{in_jk} is hidden neuron; v_{0j} is the weight bias neuron input to j ; x_i is the input neuron to i ; v_{ij} is the weight of the input neuron to hidden neuron.

By applying the calculated activation function with equation (2).

$$Z_j = f(Z_{in_j}) \quad (2)$$

Which, Z_j is the unit to j the hidden layer; Z_{in_j} is the output for the unit Z_j . For example, the activation function used is sigmoid with equation (3).

$$Y = f(x) = \frac{1}{1 + e^{-x}} \quad (3)$$

And sends this signal to all units in the output unit.

5. Each unit of output ($y_k, k = 1, \dots, m$) adds weighted input cues by using equation (4).

$$Y_{in_k} = w_{0k} + \sum_{j=1}^p z_j w_{jk} \quad (4)$$

Which Y_{in_k} is the output for the unit y_k ; w_{0k} is the bias weight for the hidden to k ; z_j is the unit to j in the hidden layer; w_{jk} is the weight of hidden neurons to output neurons by applying the activation function calculated by equation (5).

$$Y_k = (Y_{in_k}) \quad (5)$$

Which, Y_{in_k} is the output for the unit Y_k .

Stage 2: propagating backward errors

6. Each unit of output ($y_k, k = 1, \dots, m$) receives its input training pattern. Calculate the error each layer with equation (6).

$$\delta_k = (t_k - y_k) \cdot (y_{in_k}) \quad (6)$$

Where, δ_k is the weight correction factor w_{jk} ; t is a target; y_k is the neuron output to k ; y_{in_k} is output for the unit y_k . Calculate the correction for weight and bias with equation (7).

$$\Delta w_{jk} = \alpha \delta_k x_j$$

$$\Delta w_{0k} = \alpha \delta_k \quad (7)$$

Which, Δw_{jk} is the difference between $w_{jk}(t)$ and $w_{jk}(t + 1)$; Δw_{0k} is the bias weight for hidden neuron to k ; α is a

learning rate; δ_k is a correction factor weights w_{jk} ; x is the input.

7. Each hidden unit ($z_j, j = 1, \dots, p$) adds up input delta (from the units that are at the top layer) with equation (8).

$$\delta_{in_j} = \sum_{k=1}^m \delta_k w_{jk} \quad (8)$$

Which, δ_k is the weight correction factor w_{jk} ; w_{jk} is weights hidden neurons to output neurons. Count the error for each layer with equation (9)

$$\delta_j = (x_{in_j}) \quad (9)$$

Which, δ_j is the weight correction factor; v_{ij} ; δ is correction factor; x is input. Calculate corrections for weights and bias by equation (10).

$$\Delta v_{ij} = \alpha \delta_j x_i \quad (10)$$

Which, Δv_{ij} is the input neuron's weight to the neuron hidden; α is the learning rate; δ_j is a correlation factor weights v_{ij} ; x_i is the input neuron to i .

Stage 3: Fix weights and biases (Update Weight)

8. Each unit of output ($y_k, k = 1, \dots, m$) renews weights and bias ($j = 0, 1, \dots, p$) calculated with equation (11).

$$(old) = w_{jk}(new) + \Delta w_{jk} \quad (11)$$

Which, w_{jk} is the weight of neurons hidden to neurons output; Δw_{jk} ; the difference in the weight of hidden neurons to neurons the output.

Each hidden unit ($z_j, j = 1, \dots, p$) updated weight and their bias ($i = 0, 1, \dots, n$) are calculated with equation (12).

$$(new) = v_{ij}(old) + \Delta v_{ij} \quad (12)$$

Which, v_{ij} is the input neuron's weight to the neuron hidden; Δv_{ij} the difference in the input neurons' weight to the neurons hidden.

9. Test the stop conditions.

3.3 Prediction Accuracy

In this study, the Mean Absolute Percentage Error (MAPE) to evaluate errors forecasting at each period and dividing by the number the forecasting period has been used [18].

The formula for performing calculations MAPE accuracy measurement can be seen in equation (13).

$$MAPE = \frac{\sum_{t=1}^n \left| \frac{Y_t - \hat{Y}_t}{Y_t} \right|}{n} \quad (13)$$

The MAPE value is used to calculate the error value in the form of a percent. After getting the error value, it will be obtained the accuracy value of the system with the formula:

$$\text{Accuracy} : 100 \% - MAPE (\%)$$

Where, Y_t is the actual value; \hat{Y}_t predictive value; The number of periods or targets.

3.4 Manual Count

1. Preprocess Data

Preprocess Data is the process to change dataset in distance (range) 0.1 - 0.9 (Normalization) using the following equation [19]:

$$x' = \frac{0,8(x - x_{min})}{x_{max} - x_{min}} + 0,1$$

Information :

- x = Data - i
- x_{max} = The Biggest Data
- x_{min} = The Smallest Data

Table 2: Normalization Data

Months	x						y
	Field area milled (Ha)		Amount of Sugarcane Milled (Ton)		Yield (%)		
	2017	2018	2017	2018	2017	2018	
May	0,10058	0,10274	0,14783	0,31219	0,10000	0,10000	0,37632
June	0,10268	0,10513	0,37742	0,48575	0,10000	0,10000	0,54788
July	0,10951	0,11009	0,80249	0,87155	0,10000	0,10000	0,90000
August	0,10965	0,10929	0,90000	0,80259	0,10000	0,10000	0,73509
September	0,10966	0,10922	0,79430	0,65229	0,10000	0,10000	0,47707
October	0,10624	0,10095	0,47998	0,17041	0,10000	0,10000	0,10000

2. Finding the Zi value

The first step in the calculation process is input (symbolized x) to obtain the model output (symbolized by y). The process to get Y in stages begins with a search for the value Z1 (Y1).

$$Z_1 = W_{11}X_1 + W_{21}X_2 + W_{31}X_3 + b_1$$

Information :

- Z = output
- W = weight
- X = input
- b = bias

3. Find the value for Yi

After obtaining Z1, the predicted output of Y1 is obtained by applying the activation function to Z1. There are many activation functions to choose from a sigmoid activation function that looks like this:

$$\sigma(z) = \frac{1}{1 + e^{-z}}$$

Information :

- e (euler value) $\approx 2,71828182845904523536028747135$
- σ = activation of sigmoid

4. RESULT AND DISCUSSION

This section describes the test findings and analysis of the BPNN method in predicting sugarcane yields either with manual counts or Matlab training. From manual count, Y_i value obtained by activation sigmoid is the final predictive result. The next stage is an improvement in the value of weight and bias [20].

The following are the prediction results obtained from the calculation manual. Prediction results are denormalized to obtain original value [21]. The formula for denormalization data can be seen in equation (15). The results are already on denormalization shows that the prediction of sugarcane yield in years 2020 is worth 3,553,844,03066 tonnes..

$$x_{prediction} = \frac{(x'_{prediction} - 0,1)(x_{max} - x_{min})}{0,8} + x_{min}$$

Information :

- $X_{prediction}$ = Prediction Data
- x_{max} = The Biggest Data
- x_{min} = The Smallest Data

Table 3: Prediction Results of Manual Count

Month(2019)	Weight No	Weight (W)	Estimation Score (Z)	Possibility Prediction (Y)	Score Denormalization
May	W1	0,38811	(0,47211)	0,38412	401,413,29934
Juni	W2	(211,43797)	0,17510	0,54366	626,821,10436
Juli	W3	(2,60795)	1,91206	0,87125	1,089,641,46241
Agustus	W4	10,08434	1,13005	0,75585	926,601,31474
September	W5	129,78274	(0,09479)	0,47632	531,679,68808
Oktober	W6	54,08599	(2,38658)	0,08420	(22,312,83827)
	bias	0,062809678			
				Prediction Result	3.553.844,03066

Following are the results of the Matlab training with functions TRAINCG, TRAINDX and TRAINGD learning as well as fail 1000 and epoch of 2000. From table 4 below it can be seen that the predicted results are close to the results manual count is by learning function TRAINCG, where the predicted value was 3,578,326.13 and MSE worth 0.0036481 which means that the accuracy of the predictions is very high.

Table 4: Prediction Results of Matlab Training

DATA SIMULATION TEST RESULTS WITH MATLAB							TOTAL
TRAINCG	Matlab Result	0,359660	0,571860	0,885130	0,731920	0,469940	0,11422
	Denormalization	366,859	666,658	1,109,251	892,794	522,664	201,997
	MSE	0,016660	-0,023972	0,014872	0,003178	0,009134	-0,014225
TRAINDX	Matlab Result	0,414720	0,514680	0,900030	0,732620	0,433600	0,158010
	Denormalization	444,648	585,873	1,130,302	893,783	471,322	81,965
	MSE	-0,038401	0,033204	-0,000029	0,002475	0,043475	-0,058014
TRAINGD	Matlab Result	0,379070	0,544620	0,873260	0,729050	0,482230	0,258150
	Denormalization	394,282	628,173	1,092,481	888,739	540,028	223,444
	MSE	-0,002780	0,00326	0,026744	0,006041	-0,005152	-0,158150

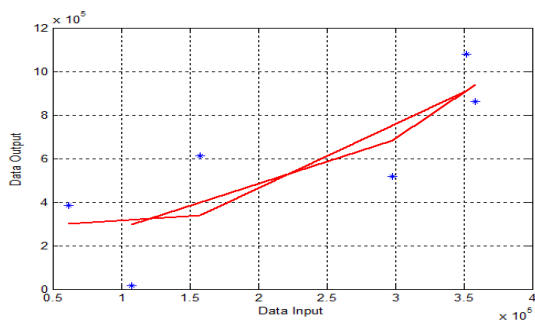


Figure1: Visualization of Matlab Prediction Results

After getting the predicted value, then we can find out the amount of error in this simulation with the MAPE formula

$$\begin{aligned} \text{MAPE} &= \text{Total Error} / \text{Lots of data (n)} \\ &= 15\% / 6 \\ &= 3\% \end{aligned}$$

The MAPE value above is used to arrive at a valuesystem accuracy with the formula: 100 % - MAPE (%).

$$\text{Accuracy} = 100\% - 3\% = 97\%$$

5. CONCLUSION

Analysis of the prediction of sugarcane production from PTPN X data using the Neural Backpropagation method Network (BPNN) has been implemented. Based on the results, Experiment with three learning functions, namely TRAINCG, TRAINGD and TRAIINGD and max file 1000 and epoch of 2000. It was concluded that the function learning that gets a predictive error rate, which is quite good with an MSE value of 0.0036481, and the prediction result of 3,578,326.13 tons is TRAINCG.

This study estimates that the predicted results of sugarcane production using weight and bias values random, and the dataset used is minimal. So that the train network process must be done continuously to get the minimum error value and the right prediction, for it is recommended for further research trying to develop the use of neural algorithms network with many datasets of classified with supervision in particular for the backpropagation model. To simplify and speed up the process, data retrieval companies can move their data to a cloud system with TOGAF and built a monitoring system long-distance using a satellite with QGIS to speed up the data retrieval process[22].

REFERENCES

- [1] FAO, "FAOSTAT () CROPS: Statistics division." [Online]. Available: <http://www.fao.org/faostat/en/#data/QC>. [Accessed: 19-Sep-2020].
- [2] W. A. Pippo and C. A. Luengo, "Sugarcane energy use:

- Accounting of feedstock energy considering current agro-industrial trends and their feasibility," *Int. J. Energy Environ. Eng.*, vol. 4, no. 1, pp. 1–13, 2013.
- [3] L. Vaddi, V. Neelisetty, B. C. Vallabhaneni, and K. B. Prakash, "Predicting crypto currency prices using machine learning and deep learning techniques," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 6603–6608, 2020.
- [4] S. A. Shinde and P. Raja Rajeswari, "A novel hybrid framework for cuff-less blood pressure estimation based on vital bio signals processing using machine learning," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 2, pp. 1556–1561, 2020.
- [5] L. S. Galvão, A. R. Formaggio, and D. A. Tisot, "Discrimination of sugarcane varieties in Southeastern Brazil with EO-1 Hyperion data," *Remote Sens. Environ.*, vol. 94, no. 4, pp. 523–534, 2005.
- [6] A. Corallo, M. Lazoi, and M. Lezzi, "Cybersecurity in the context of industry 4.0: A structured classification of critical assets and business impacts," *Comput. Ind.*, vol. 114, p. 103165, 2020.
- [7] S. P. Dangal, A. K. Das, and S. K. Paudel, "Effectiveness of management interventions on forest carbon stock in planted forests in Nepal," *J. Environ. Manage.*, vol. 196, pp. 511–517, 2017.
- [8] M. H. Saputra and H. S. Lee, "Prediction of land use and land cover changes for North Sumatra, Indonesia, using an artificial-neural-network-based cellular automaton," *Sustain.*, vol. 11, no. 11, pp. 1–16, 2019.
- [9] A. Zhu, X. Li, Z. Mo, and H. Wu, "Wind power prediction based on a convolutional neural network," *2017 Int. Conf. Circuits, Devices Syst. ICCDS 2017*, vol. 2017-Janua, pp. 131–135, 2017.
- [10] S. Jabjone and S. Wannasang, "Decision Support System Using Artificial Neural Network to Predict Rice Production in Phimai District, Thailand," *Int. J. Comput. Electr. Eng.*, vol. 6, no. 2, pp. 162–166, 2014.
- [11] E. Ismanto, N. Effendi, and E. P. Cynthia, "Implementation of Backpropagation Artificial Neural Networks to Predict Palm Oil Price Fresh Fruit Bunches," *IJISTECH (International J. Inf. Syst. Technol.)*, vol. 2, no. 1, p. 26, 2018.
- [12] C. Chen and H. Mcnairn, "A neural network integrated approach for rice crop monitoring," *Int. J. Remote Sens.*, vol. 27, no. 7, pp. 1367–1393, 2006.
- [13] S. Abdullah and E. C. Tiong, "Prediction of Palm Oil Properties using Artificial Neural Network," *J. Comput. Sci.*, vol. 8, no. 8, pp. 101–106, 2008.
- [14] I. Kopal, I. Labaj, M. Harničárová, J. Valiček, and D. Hrubý, "Prediction of the tensile response of carbon black filled rubber blends by artificial neural network," *Polymers (Basel)*, vol. 10, no. 6, 2018.
- [15] I. Prihandi *et al.*, "Implementation of Backpropagation Method for Identified Javanese Scripts," *J. Sist. dan Teknol. Inf. Indones.*, vol. 01, no. 3, pp. 0–6, 2016.
- [16] M. A. Abuzaraida and O. M. Elrajubi, "Improving laser mark detection for retinal images based on the alexnet model," *Int. J. Adv. Trends Comput. Sci. Eng.*, vol. 9, no. 4, pp. 4597–4603, 2020.
- [17] Z. Qu, W. Mao, K. Zhang, W. Zhang, and Z. Li,

- “Multi-step wind speed forecasting based on a hybrid decomposition technique and an improved back-propagation neural network,” *Renew. Energy*, vol. 133, pp. 919–929, 2019.
- [18] U. Khair, H. Fahmi, S. Al Hakim, and R. Rahim, “Forecasting Error Calculation with Mean Absolute Deviation and Mean Absolute Percentage Error,” *J. Phys. Conf. Ser.*, vol. 930, no. 1, pp. 0–6, 2017.
- [19] D. J. McCarthy, K. R. Campbell, A. T. L. Lun, and Q. F. Wills, “Scater: Pre-processing, quality control, normalization and visualization of single-cell RNA-seq data in R,” *Bioinformatics*, vol. 33, no. 8, pp. 1179–1186, 2017.
- [20] C. Edmond and A. S. Girsang, “Classification performance for credit scoring using neural network,” *Int. J. Emerg. Trends Eng. Res.*, vol. 8, no. 5, pp. 1592–1599, 2020.
- [21] Z. Liu and S. Idreos, “Main memory adaptive denormalization,” *Proc. ACM SIGMOD Int. Conf. Manag. Data*, vol. 26-June-20, pp. 2253–2254, 2016.
- [22] I. Ranggadara, N. Legowo, and Sfenrianto, “Migrating applications to clouds with TOGAF framework at MERAKI digital Indonesia,” *J. Theor. Appl. Inf. Technol.*, vol. 97, no. 9, pp. 2515–2525, 2019.