Analysis of the Artificial Neural Network Approach in the Extreme Learning Machine Method for Mining Sales Forecasting Development

Hendra Kurniawan[1], Joko Triloka*[2], Yunus Ardhan[3]

Information System-Faculty of Computer Science, Institut Informatika dan Bisnis Darmajaya, Lampung, Indonesia[1, 3]
Informatics Engineering-Faculty of Computer Science, Institut Informatika dan Bisnis Darmajaya, Lampung, Indonesia[2]

Abstract—Forecasting is an accurate indicator to support management decisions. This study aimed to mining sales forecasting on Indonesia's consumer goods companies with business warehouses engaged in the dynamic movement of large data using the Artificial Neural Network method. The sales forecasting used traditional method by inputting data and improvising simple patterns by collecting historical sales and remaining stock. Furthermore, several data variables in business warehouses were employed for sales forecasting. The study also used qualitative method to investigate the quality of data that cannot be measured quantitatively. The results showed with Mean Square Error score of 0.02716 in forecasting sales. The average accuracy generated by the Extreme Learning Machine after nine data tests is 111%. The result shows an opportunity for the company to further analyze the sales profit growth potential. The predicted value generated by Extreme Learning Machine for the last three months reaches 132%. The company's improved decision-making enlarge potential production line demonstrates the usefulness of this study.

Keywords—Artificial neural network; business warehouse; extreme learning machine; mining sales forecasting

I. INTRODUCTION

Sales forecasting is always at the forefront of decision-making and planning [1] and also for business [2]. Forecasting gives an organization the right plan to deal with future demands but does not guarantee the success of a strategy. However, its failure results of wrong decisions in marketing activities caused by the allocation of other resources which constructed on imprecise and uncertain assumptions, hence resulting to wrong decisions [3]. Sales forecasting is an important prerequisite in many aspects of sales chain management. Therefore, further optimization efforts are needed to create sales forecasts that support decision-making by the organization.

This study aimed to explore the potential of scientific forecasting and prove the optimization of existing theories. The forecasting concept uses the *Artificial Neural Network* (ANN) method, representing the human brain's performance [4]. The human brain always experiences a period of learning in the interconnected neurons. Information received by neurons is sent from one neuron layer to another [5].

This study also explained the *Extreme Learning Machine* (ELM) which is widely used in batch, sequential, and incremental learning. The method is used because of its efficient convergence and learning speed, good generalization ability, and ease of implementation [6]. Furthermore, the study built a suitable model to estimate the data generation process underlying the series. It also estimated the desired number of future observations through this model [6].

Several data variables in the business warehouse were adopted for sales forecasting using the ANN ELM method. Various studies have revealed that the ELM method has an advantage in learning speed which described in the previous section. It helps determine the effectiveness of forecasting, as well as deficiencies and necessary improvements.

II. RELATED WORK

Prianda and Widodo (2021) used ELM to project foreign tourist visits and obtained a forecast MAPE value of 7.62% [7]. Similarly, Sharma et al. [8] employed an intelligence model to predict sales using sentimental product analysis. The study made the model more productive in obtaining accurate results for each item in the product life cycle. Meanwhile, [9] applied ELM integrated with LSTM to Bitcoin price forecasting. Their study matches the different machine learning algorithms to corresponding multiscale components and constructs the ensemble prediction models based on machine learning and multiscale analysis. The results showed that the ensemble models can achieve a prediction accuracy of 95.12% with enhanced performance than the benchmark models [9]. Moreover, Cholid and Aly [10] forecasted spiral and leaf spring products for four-wheeled vehicles using the Artificial Neural Network Backpropagation method. The study employed a learning rate weight value of 0.1 of four hidden layers with error of 0.01. Several studies use various methods to predict product sales [11]–[13], compare forecasting techniques for financial prediction [14], supply chain [15]–[17], and manufacturing processes [18].

III. METHOD

This study aimed to predict sales using the ELM method, which applies qualitative approaches. It also intended to investigate, discover and explain the peculiarities of the data influence that cannot be measured or described quantitatively. Fig. 1 presents the proposed methods used in this study.

The proposed method in Fig. 1 consists of five phases. Firstly, problem formulation was formed through the question regarding implementation of forecasting with other fast and accurate methods. After finding the problem formulation, a literature study was performed by collecting sources from previous journals and articles and quarrying up information from company forecasting results.

Furthermore, the ELM method accomplished with the first stage is initializing the input weight and bias to prevent layer activation outputs from detonating or fading gradients during the normalization. Second stages are use of range between 0-1 for normalizing to ensure quality of data during training and testing phase. Training and testing stage then taking into account respectively to train model, discover and learn patterns. Testing data use to evaluate the functioning and improvement of algorithms training and optimize it for better results. The last stage of ELM is denomarlizing previous prediction results to compare the output in order to evaluate the model. The final step was to exam the *Mean Square Error* (MSE) using python pseudocode to determine the error rate of the prediction results.

For training and testing of the proposed model, the company's secondary data have been used. From the proposed model, it can be seen that the prediction result process quickly for both training and testing sets. Hence, the proposed model will provide more suitable for the task at forecasting sales compare to other techniques that described in the previous section.

A. Data Type and Source

This study used secondary data from the dataset of a retail company. The data were collected from daily sales based on the company's existing business warehouse system.

B. Data Used

The study used sales data set digitized through the company's SAP Business Warehouse system. The historical data set had 43 rows and 6 parameter features. Furthermore, a time series analysis was conducted on the previous month of sales data. This study also analyzed time, brand, and stock factors affecting sales levels through forecasting methods. The dataset design has been given in Table I.

TABLE I.   DATASET DESIGN

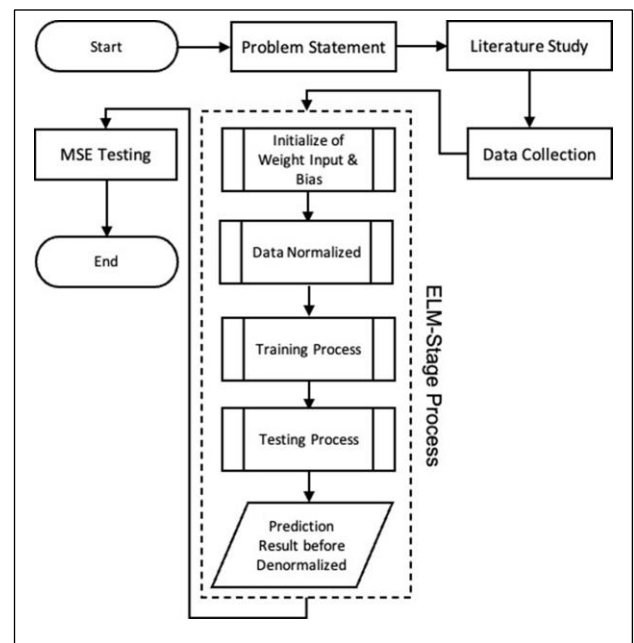| 1st Data (Time Series)	Feature Stock	Feature 1 Sales	Feature 2 Sales 2	Feature 3 Sales 3	Feature 4 Sales 4
Nov 2017	Vol	Vol	Vol	Vol	Vol
Time Series n	Vol	Vol	Vol	Vol	Vol

Fig. 1.   Proposed forecasting sales through ELM stages

C. Schematic of Research Based on Method

The ELM algorithm was used by entering historical sales data, the number of features and hidden neurons, as well as the percentage of training and testing data as initial input. Fig. 2 shows the stages outlined as follows:

a) Input weight and bias were processed randomly with values ranging between -1 to 1 based on the number of neurons then. The next step was transposing the matrix and normalizing the data using the range 0 - 1 to calculate the hidden layer output.

b) The matrix result was multiplied by the transposition before calculating the Moore-Penrose Generalized Inverse matrix using the Moore-Penrose Generalized Inverse matrix equation. The result was multiplied by the transposition of the activated hidden layer output matrix.

c) The study calculated the output results of the training process to obtain the output weight used in the testing process.

d) The testing process employed the input weight and bias obtained from the training process. The hidden layer output was then calculated using the activation function.

e) The output weight value obtained in the training process was used in the testing process to calculate the output layer as the prediction result.

f) Before the denormalization process, the study calculated the error value on all output layers not denormalized with actual data. This error value is the prediction results obtained. Additionally, the error value was tested using Mean Square Error (MSE).

g) The last step was denormalization to generate a previously denormalized value back to the original value. Fig. 2 is a flowchart of the problem-solving process with ELM.

D. Input Data

This study employed sales data detailing the monthly sales for November 2017-May 2021. The ELM method processes data by determining several criteria parameters to achieve a small error rate and optimal accuracy. Table II describes several parameters used in processing the dataset, including the input and output layers, hidden neurons, and activation function.

E. ELM Development Method Forecasting

The initial step was to find the optimal value of several input parameters tested. This step ensures that the process in the ELM method produces good predictions during training. After testing the MSE on hidden neurons, the dataset was normalized into the range of 0 – 1. This was followed by data training, testing, denormalization, and sales predictions in the following month.
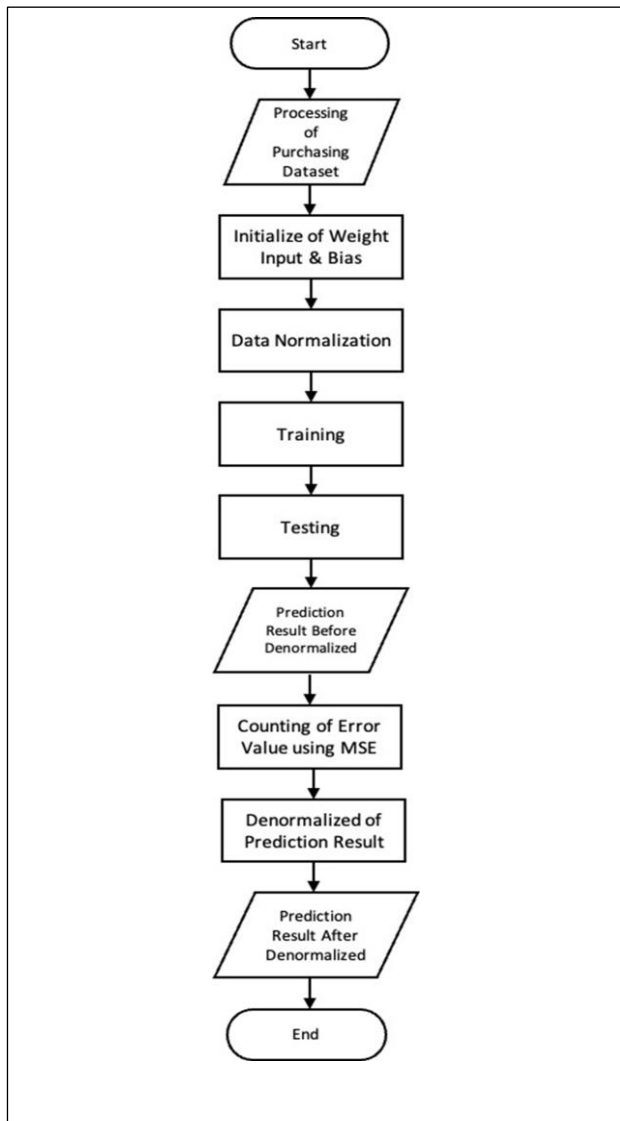

Fig. 2. ELM process diagram

TABLE II. ELM METHOD REQUIREMENT PARAMETERS

Parameter	Total	Description
Input Layer	6	Streak months, Data Stock (Label) & 4 Brand Historical Sales
Output Layer	1 Neuron	Forecasting Results
Hidden Layer	10 Testing	11 - 30 Neuron
Activation Function	1	Sigmoid Biner

IV. IMPLEMENTATION AND RESULT

A. Hidden Neuron Total Network Testing

The number of hidden neurons was tested to determine its effect on the accuracy value in implementing the ELM algorithm. The number of hidden neurons used in this test includes 11, 12, 13, 14, 15, 17, 24, 26, 27, and 30.

The tests were performed ten times, and the number of hidden neuron networks was processed into python pseudocode. Furthermore, the test was carried out repeatedly by changing the number of neurons, as shown in Table III.

TABLE III. HIDDEN NEURON TEST RESULTS

No	Total Hidden	MSE	Time (second)
1	11	0.06755	0.0515
2	12	0.02716	0.0525
3	13	0.04812	0.0528
4	14	0.08318	0.0545
5	15	0.10937	0.0526
6	17	0.12538	0.0532
7	24	0.13625	0.0550
8	26	0.21435	0.0548
9	27	0.18092	0.0545
10	30	0.24473	0.0547

Fig. 3 shows that changes occur when there is a change in the total of hidden layers. Therefore, the optimal network value was obtained in the second test with an MSE of 0.02716 and a time of 0.00525 seconds.

Fig. 3. Optimal hidden network graph plot

B. Preprocessing

Data processing was performed after creating the hidden neuron network. The data collected were then normalized using Min-Max normalization or transformed into a range of 0-1 to obtain a value for each auxiliary variable.

The input process used monthly time series, stock, and sales variables by brand. The data were divided for testing and training via Python's train_test_split function. Furthermore, the study compared the performance of the models used in forecasting sales data. The training and testing data distribution was 80% and 20%, respectively. Tables IV and V show that the output data were broken down into two parts and formed a range of numbers 0-1 equated with numbers 0-1.

TABLE IV.   DISTRIBUTION OF 80% OF TRAINING DATA AFTER NORMALIZATION

Feature 1	Feature 2	Feature3	Feature4	Feature5	Feature6
0.3552	0.3497	0.3589	0.3498	0.7480	0.7270
0.5395	0.3603	0.4630	0.3677	0.7573	0.7980
0.4781	0.3504	0.4128	0.3667	0.6723	0.7193
0.5268	0.3705	0.5047	0.3696	0.6141	0.6674
0.5049	0.3655	0.4704	0.3642	0.7085	0.5706
0.4635	0.3553	0.3978	0.3701	0.9262	0.6889
0.5078	0.3709	0.4193	0.3669	0.5677	0.6189
0.4899	0.3538	0.4004	0.3754	0.6889	0.6091
0.4777	0.3507	0.3950	0.3784	0.6787	0.6886
0.4399	0.3614	0.4304	0.3541	0.5897	0.6004
0.4511	0.3546	0.3994	0.3606	0.7527	0.6141
0.4414	0.3625	0.4619	0.3610	0.6547	0.5677
0.4758	0.3505	0.4119	0.3642	0.7157	0.6713
0.4537	0.3506	0.3798	0.3679	0.6708	0.6994
0.5057	0.3518	0.3929	0.3873	0.7114	0.6157
0.4960	0.3660	0.4483	0.3751	0.6091	0.6229
0.4642	0.3492	0.3900	0.3710	0.7454	0.6838
0.4565	0.3532	0.4274	0.3662	0.6004	0.6152
0.4644	0.3634	0.4306	0.3645	0.6813	0.6149
0.4722	0.3555	0.4084	0.3745	0.6229	0.6813
0.4781	0.3703	0.4697	0.3596	0.6935	0.7085
0.4448	0.3570	0.4504	0.3636	0.6432	0.5754
0.4375	0.3494	0.3671	0.3833	0.7270	0.7254
0.4868	0.3598	0.4206	0.3821	0.6713	0.9262
0.4590	0.3533	0.4360	0.3651	0.6152	0.6547
0.4251	0.3489	0.3942	0.3684	0.7193	0.7157
0.4282	0.3561	0.4971	0.3613	0.5754	0.6646

Feature 1	Feature 2	Feature3	Feature4	Feature5	Feature6
0.4595	0.3654	0.4490	0.3643	0.7980	0.7527
0.4700	0.3495	0.3972	0.3666	0.6838	0.6723
0.4789	0.3494	0.4008	0.3781	0.7254	0.7114
0.4622	0.3492	0.3834	0.3693	0.7294	0.7454
0.5021	0.3523	0.4390	0.3694	0.6457	0.7310
0.4964	0.3505	0.3859	0.3760	0.6886	0.7294
0.4651	0.3507	0.3809	0.3786	0.6157	0.6708

TABLE V.   DISTRIBUTION OF TESTING DATA 20% AFTER NORMALIZATION

Feature 1	Feature2	Feature3	Feature4	Feature5	Feature6
0.48451	0.35353	0.42715	0.34968	0.69842	0.72659
0.43501	0.35957	0.42615	0.34449	0.64512	0.36723
0.45851	0.34328	0.40150	0.36011	0.69533	0.67470
0.46529	0.36314	0.48058	0.46193	0.76199	0.76036
0.47845	0.34694	0.45601	0.35457	0.76622	0.79335
0.49427	0.36567	0.50819	0.35478	0.59095	0.61250
0.48471	0.35135	0.46971	0.36826	0.64059	0.67189
0.52766	0.34101	0.40406	0.37653	0.73279	0.71063
0.49897	0.35415	0.41395	0.35304	0.36723	0.78375

C. Preprocessing

The training process is conducted using 80% of the dataset. Table VI shows the data normalized and formed a training output table with Order 12 x 6. This implies 12 hidden layer networks and 6 input features.

TABLE VI.   TRAINING ORDO 12 X 6

Feature1	Feature2	Feature 3	Feature 4	Feature 5	Feature 6
0.0558	0.8430	0.6263	0.4151	0.6627	0.0252
0.0896	0.6381	0.5356	0.7703	0.3389	0.7839
0.1758	0.3225	0.5412	0.7633	0.4020	0.5811
0.4652	0.5609	0.1406	0.4446	0.7704	0.6104
0.5294	0.6224	0.4036	0.7573	0.0588	0.1303
0.6689	0.3863	0.4126	0.8024	0.8661	0.5111
0.7164	0.6305	0.5334	0.3697	0.0297	0.4240
0.7621	0.1383	0.3566	0.2204	0.9654	0.0216
0.8333	0.2910	0.8613	0.3733	0.3512	0.2307
0.8409	0.7427	0.2408	0.0436	0.0281	0.1970
0.9815	0.1281	0.4837	0.3220	0.0628	0.5891
0.9848	0.5148	0.9862	0.9852	0.2135	0.8389

Randomized training data produce output weights needed during the testing process. Table VII show the weighted output.

TABLE VII. OUTPUT WEIGHT

Output Weight (Beta)
419.442941
414.794705
713.993594
-149.313835
-473.334959
30.800049
-260.807955
-146.883101
-248.036551
-35.957942
-295.455812
32.925058

D. Testing

The testing process aimed to measure the performance of the network model built during the training process. Although the steps used were similar to the training process, all weights were taken from the training results, implying no calculation of the β weight. The data used differed from the training process explained in the preprocessing stage. The accuracy level was calculated in the same way as in the training process. The output data testing is presented in Table VIII.

TABLE VIII. OUTPUT DATA TESTING

Feature 1	Feature 2	Feature 3	Feature 4	Feature 5	Feature 6
0.4845	0.3535	0.4272	0.3497	0.6984	0.7266
0.4350	0.3596	0.4261	0.3445	0.6451	0.3672
0.4585	0.3433	0.4015	0.3601	0.6953	0.6747
0.4653	0.3631	0.4806	0.4619	0.7620	0.7604
0.4785	0.3469	0.4560	0.3546	0.7662	0.7934
0.4943	0.3657	0.5082	0.3548	0.5909	0.6125
0.4847	0.3513	0.4697	0.3683	0.6406	0.6719
0.5277	0.3410	0.4041	0.3765	0.7328	0.7106
0.4990	0.3542	0.4139	0.3530	0.3672	0.7837

Table IX shows that the results generated from the training process had an average prediction of 111%.

TABLE IX. OUTPUT TARGET TESTING VS PREDICTION TESTING RESULTS

Data To-	20% target data (Y)	Prediction Results (y)	Accuracy
Data to-1	0.4057	0.5270	130%
Data to-2	0.2980	0.5998	201%
Data to-3	0.3145	0.5494	175%
Data to-4	0.7608	0.5992	79%
Data to-5	0.4882	0.5891	121%
Data to-6	0.6600	0.5091	77%
Data to-7	0.5859	0.5569	95%
Data to-8	0.5570	0.6899	124%
Data to-9	0.4179	0.3409	82%
Average	0,4987	0,5513	111%

E. Denormalization

In the denormalization stage, the predicted test data in the 0-1 output range were converted to the actual value in kilograms. This was done to ensure that the target data and predicted results were read on a wider scale than the original value.

The denormalized data in Table X results from nine data tested and predicted to compare each accuracy per line. The accuracy from the average of the final data line was 101%.

TABLE X. DENORMALIZATION OF TARGET TESTING

Data	20% Target Data (Y)	Prediction Results (y)	%
Data to-1	57.521	67.931	118%
Data to-2	47.425	74.235	157%
Data to-3	49.077	69.851	142%
Data to-4	90.227	74.183	82%
Data to-5	64.634	73.293	113%
Data to-6	79.734	66.413	83%
Data to-7	73.014	70.497	97%
Data to-8	70.499	82.638	117%
Data to-9	58.587	51.616	88%
Average	67.367	68.250	101%
Data to-7 - 9			

F. Next Month Prediction Results

The predictions made via pseudocode python showed that the value on the 44th data or the following month from the processed dataset is 70.524 kg, as shown in Fig. 4. In Fig. 4, the red line on the y-axis explains that sales occur in the 44th data, a decrease from the previous month.
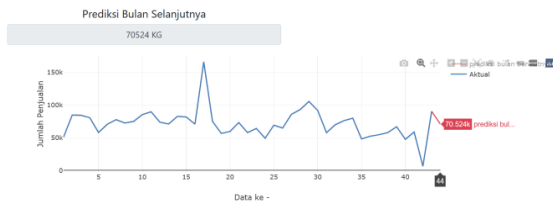
Fig. 4.   The plot of the next month's prediction results

G. Forecasting Exist Sales

Retail companies have a forecast to predict sales within the next three months. This is known through an analysis of traditional time series forecasting. Table XI and Table XII shows forecast existing and actual sales respectively. Table XIII compares the prediction made and the actual from the total forecasting, showing a gap with an accuracy of 86%.

TABLE XI.   FORECAST EXISTING

Brands	March	April	May	Total
Promina	8.996	751	15.558	8.435
SUN	36.524	3.577	33.461	24.520
Govit	1.439	122	1.962	1.174
Gowell	2.844	1.233	21.270	8.449
Total	49.802	5.682	72.251	42.579

TABLE XII.   ACTUAL SALES

Brands	March	April	May	Total
Promina	9.697	904	22.218	10.940
SUN	36.771	3.708	54.311	31.597
Govit	1.108	64	1.348	840
Gowell	11.010	1.553	12.349	8.304
Total	58.587	6.229	90.227	51.681

TABLE XIII.   COMPARISON OF EXIST VS ACTUAL FORECASTING

Brands	March	April	May	*Average*
Forecast	49.802	5.682	72.251	42.579
Actual	58.587	6.229	90.227	51.681
Accuracy	85%	91%	80%	82%

Where:

Forecast: The forecasting performed by Retail companies in units of volume

Actual: Sales in March, April, and May

Accuracy: *Forecast*/Actual Sales

The average data for three months of prediction in Table XIII confirm that the average accuracy is only 82%. A less-than-actual plot accuracy warns the company of insufficient stock available to supply future sales. This means that sales would not be maximized.

H. Comparison of Exist vs ELM Development Methods

ELM exist vs development method matrix: From the implementation of sales forecasting proposed model, a comparison matrix could be made between the Exist versus ELM Method as given in Table XIV. Based on the forecast matrix, it can be seen that the average 3-month forecast value reached 68,250 kilograms by using the ELM method, while the traditional method reached 42,579 kilograms. The average value of 3 months is actual sales 51,681 kilogram has been used for comparison the ELM and traditional method. Based on results, it can be concluded that predictions by ELM provide more better accuracy to provide adequate stock for the next sale. It mean that sales be able to run optimally for company. Meanwhile, based on accuracy matrix the ELM method can achieve admirable accuracy values, this is due to the stock feature as basic calculation label are carried out to acquire the smallest error value as well as several tests also has been done. Some of these features are support in providing stimulation to achieve better accuracy.

V.   CONCLUSIONS

Based on the test results and discussion, the following conclusions were obtained:

a) As measured by the Mean Square Error (MSE) in forecasting sales results, the error rate is 0.02716 with a time of 0.0525.

b) The ten neuron network tests showed that more hidden neurons do not measure algorithm optimization. Therefore, using 12 of 30 hidden neurons produces fewer error values.

c) The average accuracy value generated by the ELM forecasting method when testing nine data is 111%. This illustrates an opportunity for the company to further analyze the sales profit growth potential.

d) The predicted value produced by ELM for the last three months reaches 132% compared to traditional methods, which only achieve 82%.

TABLE XIV. COMPARISON OF THE EXIST VS ELM METHOD

Matrix	Method	Average Unit (Kg)	Description
Forecast	Exist (Traditional)	42.579	The average 3-month forecast value conducted by the company using the traditional method is 42.579 kilograms. This value is obtained from Table XI.
	ELM	68.250	The average value of the 3-month forecast using the ELM method is 68.250 kilograms. This value is obtained from Table XIII.
	Actual Average 3 Months	51.681	The average value of 3 months compared to actual sales is 51.681 kilograms. This value is obtained from Table XIII.
Accuracy	Exist (Traditional)	82%	Comparison of Exist / Traditional Forecast Values compared to Actual. Value calculated using spread sheet gets a relatively good accuracy value but has not yet reached the optimal value. This is because the calculated forecast still applies manual estimation as a measuring basis by dividing average monthly sales and adding sales estimates according to forecast growth.
	ELM	132%	Comparison of ELM Forecast Value compared to Actual. In line with the proof method described in the previous discussion, the accuracy value of the ELM method could achieve a very good accuracy value. This is due to the featured stock, the reference label for calculations, and several tests are conducted to get the smallest error value and the best accuracy. Some of these features help in providing stimulation to achieve optimal accuracy.

Sales forecasting still requires extensive machine learning and statistics knowledge. Addition features of input layer such as number of workers, demographic trends, and behavioral indicators should be involved in future studies to improve the model.

ACKNOWLEDGMENT

REFERENCES

[1] F. Petropoulos et al., "Review forecasting: theory and practice," Int J Forecast, vol. 38, no. 3, pp. 705–871, 2022.

[2] O'Trakoun, J. "Business forecasting during the pandemic". Bus Econ vol. 57, pp. 95–110, 2022.

[3] A. K. K. P. Dewi, "Penerapan metode extreme learning machine pada peramalan jumlah penumpang penerbangan domestik di Kepulauan Riau," in Prosiding Statistika, pp. 589–596, 2021.

[4] A.G. Bonab, "A comparative study of demand forecasting based on machine learning methods with time series approach," Journal of Applied Research on Industrial Engineering, vol. 9, no. 3, pp. 331–353, 2022.

[5] M. S. R. Dastres, "Artificial neural network systems," International Journal of Imaging and Robotics (IJIR), vol. 21, pp. 13–25, 2021.

[6] N. A., R. M. J., M. and A. Z. B.P. Ooi, "A study of extreme learning machine on small sample-sized classification problems," J. Physics.: Conference Series, vol. 2107, pp. 1–7, 2021.

[7] B. Prianda and E. Widodo, "Perbandingan metode seasonal ARIMA dan extreme learning machine pada peramalan jumlah wisatawan mancanegara ke Bali," Barekeng: J. Math. & App, vol. 15, pp. 639–650, 2021.

[8] N. G. J. R. and M. B. A. K. Sharma, "An Intelligent model for predicting the sales of a product," in 2020 10th International Conference on Cloud Computing, Data Science & Engineering (Confluence), pp. 341–345, 2020

[9] C. Luo, L. Pan, B. Chen and H. Xu. "Bitcoin price forecasting: an integrated approach using hybrid LSTM-ELM models". Mathematical Problems in Engineering, vol. 2022, 2022.

[10] D. A. F. Cholid, "Implementation of product sales forecast using artificial neural network method," International Journal of Information System and Technology (IJISTECH), vol. 5, pp. 153–162, 2021.

[11] A. U. P. Paduloh, "Analysis and comparing forecasting results using time series method to predict sales demand on Covid-19 pandemic era," Journal of Engineering and Management in Industrial Systems, vol. 10, pp. 37–49, 2022.

[12] P. Smirnov & S. Vladimir. "Forecasting new product demand using machine learning". Journal of Physics: Conference Series. 1925. (1): 012033, 2021

[13] D. H. and R. A. N. V. Wineka, "Sales Forecasting by iusing exponential smoothing method and trend method to optimize product sales in PT. Zamrud Bumi Indonesia during the Covid-19 pandemic," Int. J. Eng. Sci. Inf. Technology, vol. 1, pp. 59–64, 2021.

[14] M. U. G. and A. M. O. O. B. Sezer, "Financial time series forecasting with deep learning: A systematic literature review: 2005–2019," Appl. Soft Comput. J, vol. 90, pp. 2005–2019, 2020.

[15] E. Beh., G. T. and R. G. M. Abolghasemi., "Demand forecasting in supply chain: the impact of demand volatility in the presence of promotion," Comput Ind Eng, vol. 142, 2020.

[16] S. S. W. S. and M. M. N. E. Fradinata, "Compare the forecasting method of artificial neural network and support vector regression model to measure the bullwhip effect in supply chain," J. Mech. Eng. Sci, vol. 13, pp. 4816–4834, 2019.

[17] M. Alnahhal, D. Ahrens, and B. Salah, "Dynamic lead-time forecasting using machine learning in a make-to-order supply chain," Applied Sciences, vol. 11, no. 21, p. 10105, Oct. 2021.

[18] R. Kannan., H. A. Abdul Halim., K. Ramakrishnan, et al. Machine learning approach for predicting production delays: a quarry company case study. J Big Data 9, no. 94, 2022.