Journal of Applied Informatics and Computing (JAIC)
Vol.9, No.5, October 2025, pp. 2461~2467
e-ISSN: 2548-6861

Extreme Learning Machine Method Application to Forecasting Coffee Beverage Sales

Yusuf Sutanto [1]*, Heribertus Ary Setyadi [2], Wawan Nugroho [3]**, Budi Al Amin [4]*****
* Informatics, Universitas Dharma AUB
** Information System, Universitas Bina Sarana Informatika
*** Accounting Information System, Universitas Bina Sarana Informatika
yusuf.sutanto@stie-aub.ac.id [1], heribertus.hbs@bsi.ac.id [2], wawan.wgh@bsi.ac.id [3], budi.bdm@bsi.ac.id [4]

Article Info

Article history:

Received 2025-07-25
Revised 2025-09-11
Accepted 2025-09-20

Keyword:

Coffee Baverage,
Extreme Learning Machine,
Forecasting,
Sales.

ABSTRACT

Sales estimates can be used to set product prices and increase expected profits. Flyover coffee shop Karanganyar does not have a methodical forecasting method to estimate and predict their need/demand for coffee beverage products. Two previous research that used Extreme Learning Machine (ELM) method in other predictions stated that ELM method has high accuracy and fast compilation time. Another research predicted jeans sales using the ARIMA model and produced an accuracy of 17.05% based on the MAPE (Mean Absolute Percentage Error) method. Menstrual cycle prediction using the Long Short-Term Memory (LSTM) method produces a MAPE value of 7.5%. Two advantages of ELM method from two previous research were used as the basis for selecting ELM method used in our study. To help predict sales of coffee beverage menus, this research utilized an artificial neural network method using ELM algorithm. ELM method consists of an input layer and an output layer connected through a hidden layer. Data used for the test was daily sales data for a month. Data used for this study consisted of 215 data samples. Daily sales data at the Flyover coffee shop were collected from June to December 2024. Based on the results and analysis of error values using MAPE method, an average error value was 8.274%. From comparison of original data results and prediction data, an average MAPE error value the best number of features and hidden neurons is 5.65%.

I. INTRODUCTION

Indonesia is one of the best coffee producer countries in Asia. There are four types of coffee known to Indonesian people, include Arabica coffee, Robusta coffee, Liberica coffee, and Excelsa coffee [1]. The most well-known coffees are Arabica coffee and Robusta coffee because they are known to have economic value in coffee trading market [2]. Coffee producing areas that are famous for producing coffee include Java, Jambi, Lampung, Aceh, Flores, and Toraja [3]. Coffee is one of plantation commodities that has a high economic value among other plantation crops, and plays an important role as a source of foreign exchange for countries [4]. According from the International Coffee Organization (ICO) data , in April 2024 Indonesia was ranked 2nd as a coffee producer after Vietnam in Asia Pacific region [5].

For Indonesian people, sit around while sipping coffee is a tradition to celebrate togetherness values and strengthen brotherhood between people [6]. Sipping coffee, started as a tradition, and then branch out into a business opportunity by opening a coffee shop [7]. The culture among young people who like to hang out has triggered coffee shop growth businesses in Indonesia, especially in Surakarta city. As time goes by, more and more Indonesian people are motivated to open a coffee shops business, making it easier to find coffee shops [8]. As a coffee producing country, Indonesia is ranked 2nd in the world for consuming the most coffee [9].

Selling estimates are made to plan production, purchasing raw materials and other material purchases. Sales estimates can be used to set product prices and increase expected profits [10]. Sortilege is an activity that aims to predict an event in the future and is a tool for planning [11]. A prediction process can help companies in making corporate decisions such as determining optimal inventory levels and assisting in stock management to reduce planning errors risk to a minimum [12]. Sales forecasting or prediction can be done by building a calculation model based on previous sales data. Currently, Flyover coffee shop Karanganyar does not have a methodical forecasting method to estimate and predict their need/demand for coffee beverage products. Forecasting is only done by relying on previous period demand or using intuition [13]. This may cause an operational management process not running efficiently and effectively. Therefore, in order for sales estimates to be accurate, sales forecasting needs to be carried out based on structured analysis [14]. To help predict sales of coffee drink menus, this research utilized an Artificial Neural Network method using Extreme Learning Machine (ELM) algorithm [15].

Artificial Neural Networks (ANN) are a subdivision of machine learning. ANN is a technique used to build intelligent programs with modeling that simulates how neural networks work in human brain [16]. ELM is a feed-forward artificial neural network method that only has one hidden layer or is commonly known as a single-hidden layer feed-forward neural network. An ELM learning method was created to overcome feedforward artificial neural networks weaknesses, especially in terms of learning speed [17]. This method uses a neural network algorithm with one hidden layer to calculate an output weights using least squares method. In learning using conventional gradient based learning algorithms such as backpropagation (BP), all parameters in feedforward ANN including input weight and hidden bias must be determined manually [18]. These parameters are also interconnected between one layer and another, so they require a long learning speed and often get stuck in local minima. In ELM, parameters such as input weight and hidden bias are chosen randomly, with the result ELM has a fast learning speed and is able to produce good generalization performance [19]. The ELM method has a simpler and more effective mathematical model than feedforward artificial neural networks [20].

Previously, several studies have been made on predicting sales theme and ELM. They include research by Nariyana that predicting number of coffee shop purchases in Surabaya using catboost with leave-one-out cross validation. This study uses a quantitative approach and novel solutions to develop a machine learning-based purchase prediction application and strategies to increase purchase volume for three coffee shops in Surabaya. The Catboost model generate an MAE of 0.91 and a MAPE of 15%, surpass LightGBM's MAE of 1.13 and a MAPE of 18%. These results indicate CatBoost effectiveness for coffee shop industry with good accuracy. CatBoost can predict purchase volume accurately, and help identify target markets [21].

Research by Setiadi comparing ELM and Support Vector Regression SVR methods to predict stock prices in four biggest Indonesian banks. The dataset used comes from Yahoo Finance site because those data displayed is real data that describes the latest stock movements. The method evaluation is carried out by comparing Mean Absolute Percentage Error (MAPE) values generated from two prediction methods. Based on evaluation results, ELM and SVR methods are effective in predicting holdings prices of four biggest Indonesian banks. From computation time results, ELM method is more efficient with an average of 0.006 seconds, compared to the SVR method with an average computation time of 0.694 seconds [22].

Research by Das purpose to analyze the Wholesale Price Index (WPI) of India and present a suitable forecasting strategy for this index. This reseacrh uses ELM approach to estimate WPI. WPI clustering reveals that the sixty individual food items under category of ‘food product manufacturing’ have varying trends, patterns and characteristics. The proposed ELM achieves the maximum number of high accuracy cases (almost 87%) among all the approaches used. It also outperforms other approaches in terms of the maximum number of indices with respect to the comparison of forecast MAPE and forecast RMSE. This study concludes that the proposed ELM is a suitable prospect to provide effective forecasts of these sixty indices [23].

Permata and Habibi produce a study to predict jeans sales using Autoregressive Integrated Moving Average (ARIMA) model. They used sales data from 2018 to 2022. The model values obtained in this study were ARIMA(3, 1, 1) and ARIMA(2, 1, 0). MAPE method was used to evaluate forecasting accuracy and produced a value of 17.05% [24]. Other research on menstrual cycle prediction using the Long Short-Term Memory (LSTM) method produced a MAPE value of 7.5% [25].

Previous research that also predicted coffee beverage sales used Catboost model while our study used ELM method. Two previous research that used ELM method in other predictions stated that ELM method has high accuracy and fast compilation time. Two advantages of ELM method from two previous research were used as the basis for selecting ELM method used in our study.

II. METHOD

Data collection used in this study was obtained from direct interviews with the owner of flyover coffee shop in Karanganyar by requesting sales data from June 2024 to December period 2024 as primary data.

A system network design used in coffee beverage sales prediction system process uses the implementation of ANN development model, that is ELM method. This method is an innovative model using the Single-hidden Layer FeedForward Networks (SLFNs) algorithm. SLFNs have an input layer architecture, a single hidden layer, and an output layer. An input layer is used to process input data in the form of features such as X1, X2, X3, X4, and added target variables (T). Each input layer neuron formed is connected to a hidden layer neuron which is connected by weight and bias variables with randomly obtained values. The ELM method network has its own architecture which is presented in Figure 1.

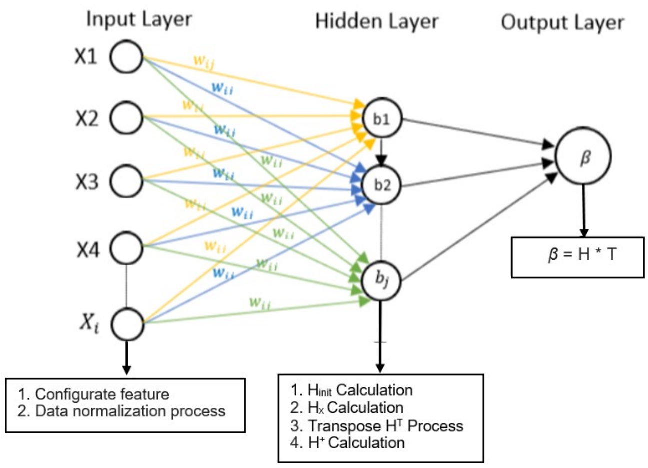

Figure 1. ELM method network architecture

Figure 1 explains that ELM method consists of an input layer and an output layer connected through a hidden layer. In addition to inputting a dataset to input layer, several processes are performed, including preprocessing. Preprocessing is an initial stage in processing raw data in data mining to facilitate its processing by system. Preprocessing in ELM method involves feature formation and data normalization. Feature establishment involves grouping data according to input features count. The grouping of input data is always accompanied by a target value. Continued with the data normalization process using the following equation:

$$V' = \frac{v - min}{max - min} \quad (1)$$

Process that occurs in hidden layer aims to obtain a pattern or formula for ELM learning method output , named output weight (β). Process steps for obtaining an output weight are as follows:

1. After the input is determined by several features, the next step is initialize an input weight (w) and bias (b) values randomly with values between -1 and +1 or 0 and 1. These grade of weight and bias are values that connect an input value with an output value of hidden neuron.
2. Hidden neuron output value (H_{init}) calculation is determined by the previously randomly decided weight and bias values. An equation in H_{init} calculation process is as follows:
$$H_{init} = \left(\sum_{k=1}^{n} W_{jk} . X_{ik}\right) + b_j \quad (2)$$
3. Activation function calculating from Hinit results. An activation function output value is denoted $H(x)$ and use binary sigmoid activation function using following equation:
$$H_{(x)} = \frac{1}{1 + e^{-H_{init}}} \quad (3)$$
4. Transpose the hidden layer output matrix (H_{init}) with the sigmoid activation function ($H_{(x)}$). This transposed output is denoted H^T [26].
5. Multiplication between transpose result matrix (H^T) and an output matrix of hidden layer with sigmoid activation function ($H_{(x)}$).
6. The matrix inverse process on result of multiplying matrices H^T and $H_{(x)}$ uses an Elementary Row Operation (OBE) method. OBE method is carried out by replacing matrix elements with identity matrix elements to produce a new identity matrix. This is continued by multiplying an inverse result value with an output value of hidden layer that has been transposed using an equation:
$$H^+ = \left(H^T . H_{(x)}\right)^{-1} . H^T \quad (4)$$
7. The final step is to obtain ELM method output value by calculating an output weight (β) using a value from Moore-Penrose generalized inverse process multiplied by target value
$$\beta = H^+ . T \quad (5)$$

[27]

Accuracy evaluation uses MAPE method to calculate an absolute average error percentage. Table I shows MAPE values range.

TABLE I
MAPE RANGE

MAPE value	Prediction Accuracy
< 10%	Very accurate
10% - 20%	Accurate
20% - 50%	Quite accurate
> 50%	Inaccurate

[28]

III. RESULT AND DISCUSSION

A. Data Trial

The data used as test material is data on the number of daily sales for a month and will be used as input data. In the testing conducted, the data used for this study consisted of 215 data samples. The daily sales data at Flyover coffeesshop was from June to December 2024. The data is used as a dataset as presented in Table II.

TABLE II
SALES DATA FOR JUNE-DECEMBER 2024

No	Date	Sum
1	06/01/2024	53
2	06/01/2024	38
3	06/01/2024	45
4	06/01/2024	71
5	06/01/2024	77
6	06/01/2024	27
7	06/01/2024	73
8	06/01/2024	51
…	…	…
215	31/12/2024	72

The data used as the dataset is only data in the sum column. Feature formation is an input layer, symbolized by X and always accompanied by a target value. The data in the sum column will be processed in feature formation, with the results presented in Table III.

TABLE III
FEATURE FORMATION

No	X1	X2	X3	X4	Target
1	52	39	44	72	77
2	37	46	70	78	27
3	44	72	76	28	73
4	70	78	26	74	51
5	76	28	72	52	59
6	26	74	50	60	63
7	72	52	58	64	71
8	50	60	62	72	90
…	…	…	…	…	…
211	37	37	40	64	72

Feature X1 is sum of sales on the first day of June, X2 is sum of sales on the second day of June, and so on for X3 and X4. Data in Table III shows 211 data sets, because the first target uses input data (features) from the previous four sets (according to count of initialized features). Therefore, the first target data is from June 5, 2024, and the data from June 1-4, 2024, is used for input data X1-X4.

Data in Table III is secondary data that will be further processed through data normalization process using min-max normalization. Before performing data normalization, the data is first divided into training and testing data in a 90%:10% ratio. This data division is performed before normalization process to obtain optimal error values. From data division results, minimum and maximum values for training data and testing data are obtained as follows: training data: min=20 and max=111, testing data: min=34 and max=3. The next step is to calculate normalization matrix for training data and testing data using equation (1).

TABLE IV
TRAINING DATA NORMALIZATION

No	X1	X2	X3	X4	Target
1	0,3483	0,2022	0,2584	0,573	0,6292
2	0,1798	0,2809	0,5506	0,6404	0,0674
3	0,2584	0,573	0,618	0,0787	0,5843
4	0,5506	0,6404	0,0562	0,5955	0,3371
5	0,618	0,0787	0,573	0,3483	0,427
6	0,0562	0,5955	0,3258	0,4382	0,4719
7	0,573	0,3483	0,4157	0,4831	0,5618
8	0,3258	0,4382	0,4607	0,573	0,7753
…	…	…	…	…	…
190	0,0787	0,0899	0,2247	0,4157	0,4944

TABLE V
TESTING DATA NORMALIZATION

No	X1	X2	X3	X4	Target
191	0,2542	0,1186	0,2203	0,5763	0,678
192	0,1356	0,2881	0,4576	0,678	0,7288
193	0,1695	0,6441	0,7119	0,7797	0,661
194	0,6102	0,661	0,7458	0,7966	0,5932
195	0,7119	0,1017	0,6271	0,3559	0,4576
196	0,1864	0,5763	0,3559	0,5254	0,4915
197	0,6441	0,3051	0,4068	0,5085	0,6271
198	0,2712	0,4407	0,4746	0,6441	0,9492
…	…	…	…	…	…
211	0,0508	0,0508	0,1017	0,5085	0,6441

Tables IV and V presented data normalization process results, which range from 0 to 1. The training and testing data were divided according to requirements so In this study, used a 90:10 data division. Continue processing training data with input weight value or initial weight value and bias obtained from input weight initialization and bias values randomly with values between 0 and 1. Initial weight and bias values are presented in Table VI.

TABLE VI
INITIAL WEIGHT AND BIAS

No	W1	W 2	W 3
1	0.26	0.23	0.72
2	0.85	0.35	0.49
3	0.18	0.76	0.03
4	0.43	0.22	0.82
No	**B1**	**B 2**	**B 3**
1	0.3	0.17	0.43

Hidden layer output value calculation using H_{int} function according to Equation 2. Hint calculation results are presented in Table VII.

TABLE VII
H_{INT} FUNCTION RESULTS

No	H_{int1}	H_{int2}	H_{int3}
1	0,6911	1,1257	0,8369
2	0,7598	1,2314	0,8628
3	0,5103	1,1199	0,9396
4	0,8725	1,5958	1,1011
5	0,6582	1,1678	1,0982
6	0,6311	1,1597	0,7411
…	…	…	…
190	0,9014	1,5469	1,1432

Calculation of binary sigmoid activation function value of the hidden layer output using the H_x function according to Equation 3. Activation function calculation results are presented in Table VIII.

TABLE VIII
BINARY SIGMOID ACTIVATION FUNCTION RESULTS

No	H_{x1}	H_{x2}	H_{x3}
1	0,6662	0,755	0,6978
2	0,6813	0,7741	0,7032
3	0,6249	0,754	0,719
4	0,7053	0,8314	0,7505
5	0,6589	0,7627	0,7499
6	0,6527	0,7613	0,6772
…			
190	0,7112	0,8245	0,7583

Output activation function matrix transpose of the hidden layer, the result of which is denoted by H^T and presented in Table IX.

TABLE IX
TRANSPOSE RESULT

Symbol	1	2	3	4	…	190
H_{x1}	0,6662	0,6813	0,6249	0,7053	…	0,7112
H_{x2}	0,755	0.7741	0,754	0,8314	…	0,8245
H_{x3}	0,6978	0,7032	0,719	0,7505	…	0,7583

Calculate multiplication of activation function data H_x with transpose data H^T. Nultiplication of H_x and H^T results are presented in matrix below.

$$\begin{bmatrix} 108,968 & 122,983 & 115,013 \\ 123,162 & 139,106 & 130,002 \\ 115,014 & 128,897 & 120,184 \end{bmatrix}$$

Determining inverse value of the product of transpose value and binary sigmoid activation function and calculation results are shown in matrix below.

$$\begin{bmatrix} 2.991 & -3.4324 & 0.8505 \\ -11.633 & 10.233 & 0.0641 \\ 9.6144 & -7.6898 & -0.8744 \end{bmatrix}$$

Calculating product of inverse and transpose values, this result is called the Moore-Penrose Generalized Inverse (H^+)

TABLE X
GENERALIZED INVERSE MOORE-PENROSE RESULT

Symbol	1	2	3	4	…	190
H_{x1}	-0,082	-0,206	0,011	-0,348	…	-0,174
H_{x2}	0,073	0,174	0,041	0,316	…	0,164
H_{x3}	0,005	0,014	-0,049	-0,003	…	-0,005

The final process is calculating output weight value according to Equation (5) and the result is:

$$\begin{bmatrix} 0.7501 \\ 0.3159 \\ 0.6152 \end{bmatrix}$$

Output weight values obtained in the previous process will be used in the testing process to obtain predicted values. The testing process also requires normalized test data, result of which are shown in Table XI.

TABLE XI
H_{INIT} RESULTS OF TESTING PROCESS

No	H_{init1}	H_{init2}	H_{init3}
191	1,6001	2,1985	1,5559
192	1,6498	2,1113	1,6638
193	1,6211	2,1024	1,6597
194	1,6012	2,1021	1,5654
195	1,5987	1,9897	1,6301
…	…	…	…
211	0,9213	1,3789	0,9605

The next step is to find activation function value from testing data and the results are presented in Table XII.

TABLE XII
RESULTS OF ACTIVATION FUNCTION OF THE TESTING PROCESS

No	H_{init1}	H_{init2}	H_{init3}
191	0,8301	0,8897	0,8261
192	0,8402	0,8931	0,8389
193	0,8341	0,8902	0,8397
194	0,8303	0,8911	0,8302
195	0,8318	0,8839	0,8361
…	…	…	…
211	0,7146	0,8001	0,7225

Next, calculate predicted value (output layer) using an equation:

$$Y = H.\beta \tag{6}$$

TABLE XIII
OUTPUT LAYER RESULTS (PREDICTION)

No	Result
1	0.4222
2	0.6255
3	0.2821
4	0.7645
5	0.3256
6	0.5895
7	0.2344
8	0.3591
9	0.5104
…	…
21	0.6922

The output layer results are converted to an original data by calculating data denormalization using the equation:

$$v = v'(max - \min) + min \tag{7}$$

Data denormalization process results are shown in Table XIV.

TABLE XIV
DATA DENORMALIZATION RESULT

No	Result
1	58.9094
2	70.8945
3	50.5254
4	80.4622
5	56.7111
6	69.7518
7	50.7703
8	57.3787
9	69.0475
…	…
21	77.5807

B. Results and Analysis of Error Values Using MAPE Method

MAPE is a measurement method to evaluate error value compared to predicted results calculated using the equation:

$$MAPE = \sum_{t=1}^{n} \frac{x_t - y_t}{x_t} \, x \, 100\% \tag{8}$$

TABLE XV
ERROR VALUE EVALUATION RESULTS

No	Error Value
1	6.493%
2	7.929%
3	9.776%
4	13.481%
5	9.983%
6	5.741%
…	…
21	7,751%

In calculating error value of all entire data, average error value produced was 8.274%.

C. Results And Trials Of Combination Of Data Features Count, Number Of Hidden Neurons, And Number Of Epochs.

Based on test results, the best and most suitable data features count was determined to produce the lowest error value in coffee sales prediction system. The feature determination used the previous day's sales figures for each feature created. Results of testing variations in data features count and the number of hidden neurons are shown in the average MAPE value in percent in table XVI.

TABLE XVI
THE AVERAGE MAPE VALUE OF VARIATIONS COMBINATION IN DATA FEATURES COUNT, HIDDEN NEURONS COUNT, AND EPOCHS COUNT

Fitur Count	Hidden Neurons Count (%)													
	2		3		4		5		6		7		8	
	Error	Ep	Error	Ep	Error	Ep	Error	Ep	Error	Ep	Error	Ep	Error	Ep
1	6,36	2	9,62	3	9,59	4	15,12	5	17,28	6	12,44	7	12,54	8
2	9,62	4	9,32	6	7,67	8	17,23	10	13,83	12	28,18	14	26,16	16
3	9,54	6	6,80	9	11,01	12	15,35	15	27,32	18	15,13	21	22,24	24
4	9,54	16	8,40	12	8,28	16	17,35	20	11,35	24	17,68	28	11,23	32
5	10,23	10	6,32	15	14,53	20	17,45	25	13,23	30	21,35	35	27,45	40
6	11,23	12	5,65	18	7,87	24	13,23	30	17,44	36	14,88	42	21,23	48
7	13,32	14	9,02	21	6,31	28	14,13	35	12,32	42	24,66	49	12,25	54
8	7,21	16	8,50	24	18,45	32	16,34	40	19,45	48	12,45	54	22,86	64

Table XV explains that the best value is found in table in row 6, representing best data features count. Column 3 representing the best hidden neurons count and has been epoched 18 times. This is because this value has the smallest MAPE error, at 5.65%.

D. Results and Trials Discussion of Combinations of Variations Data Features Count, Hidden Neurons Count and the Number of Epochs

From Table XV, it can be concluded that features count, hidden neurons count, and the number of epochs affect the prediction results obtained. It also shows that more hidden neurons, the greater resulting error value. This is due to overfitting, which is the formation of too many patterns with a small (simple) dataset, making them difficult to identify, ultimately leading to a significant increase in error.

Based on test results, this can be used as a reference in finding the best prediction value, or the one closest to an original value, and producing the lowest MAPE error in this study. Therefore, this research will test sales data as predictive data to find the best predictive value using the best features count and hidden neurons, namely X6 and 3, with 18 epochs. This test results can be seen in Table XVII.

TABLE XVII
COMPARISON OF ORIGINAL DATA AND PREDICTED DATA

Data	Original Value	Prediction Value	MAPE (%)
1	62	60,91	3,32
2	78	70,89	7,93
3	55	50,53	9,78
4	94	90,46	2,73
5	62	60,71	3,63
6	75	70,75	4,39
7	42	40,77	5,19
8	65	60,38	5,66
9	76	70,05	9,03
10	52	49,69	2,56
11	33	35,89	5,57
12	86	78,99	7,07
13	61	59,93	3,34
14	55	49,83	7,72
15	54	59,32	5,93
16	39	36,12	4,94
17	35	34,70	3,61
18	42	43,56	6,24
19	62	67,19	6,65
20	73	77,58	7,75

Table XVI shows prediction results using the best features count is 6 and the best hidden neurons count is 3. From the results obtained, the average value of MAPE error value from the best number of features and hidden neurons is 5.65%.

IV. Conclusion

ELM method used in this research on predicting coffee beverage sales produced a small error value of 8.274% for the 4th number of features and the 3rd number of hidden neurons. The experiment results that has been carried out 8 times using a data ratio of 90%:10% for training data and testing data, using binary sigmoid activation function in searching for hidden layers count. The best features obtained the best hidden layers count is 3, the best features count is 6, and the number of epochs stops at 18. Obtained the error rate tested by MAPE calculation is 5.65%. Hidden layers count affects the calculation results, and in input layer there is a feature formation whose number can also affect the calculation results.

References

[1] A. S. R. M. Sinaga, R. E. Putra, and A. S. Girsang, "Prediction Measuring Local Coffee Production And Marketing Relationships Coffee With Big Data Analysis Support," *Bull. Electr. Eng. Informatics*, vol. 11, no. 5, pp. 2764–2772, 2022, doi: 10.11591/eei.v11i5.4082.

[2] F. Orduna-Cabrera *et al.*, "Short-Term Forecasting Arabica Coffee Cherry Yields by Seq2Seq over LSTM for Smallholder Farmers," *Sustain.*, vol. 17, no. 9, pp. 1–14, 2025, doi: 10.3390/su17093888.

[3] L. Setiyani and W. H. Utomo, "Arabica Coffee Price Prediction Using the Long Short Term Memory Network (LSTM) Algorithm," *Sci. J. Informatics*, vol. 10, no. 3, pp. 287–296, 2023, doi: 10.15294/sji.v10i3.44162.

[4] J. Zhao, B. Zong, and L. Wu, "Site Selection Prediction for Coffee Shops Based on Multi-Source Space Data Using Machine Learning Techniques," *ISPRS Int. J. Geo-Information*, vol. 12, no. 8, pp. 1–24, 2023, doi: 10.3390/ijgi12080329.

[5] S. Meeprom and A. Kokkhangplu, "Customer experience and satisfaction in coffee consumption: an experiential marketing perspective," *Cogent Bus. Manag.*, vol. 12, no. 1, p., 2025, doi: 10.1080/23311975.2025.2450296.

[6] W. Wahyuningsih and P. T. Prasetyaningrum, "Enhancing Sales Determination for Coffee Shop Packages through Associated Data Mining: Leveraging the FP-Growth Algorithm," *J. Inf. Syst. Informatics*, vol. 5, no. 2, pp. 758–770, 2023, doi: 10.51519/journalisi.v5i2.500.

[7] Y. L. Chen and Y. C. Chen, "Impact of Brand and Testing Method on Coffee Taste Perception between Genders: A Comparative Study of Two Leading Coffee Brands in Taiwan," *SAGE Open*, vol. 15, no. 1, pp. 1–12, 2025, doi: 10.1177/21582440251316087.

[8] R. S. Hendiarto, R. F. Laksana, and F. Alhafizh, "Influence of Store Atmosphere and Price on Costumer Loyality (Study at Diantara Kopi Coffee Shop in Bandung)," *Owner*, vol. 8, no. 3, pp. 2834–2842, 2024, doi: 10.33395/owner.v8i3.2088.

[9] A. W. Gunawan, A. W. Muhaimin, and R. I. Sitawati, "High Quality Product, Good Services, and Competitive Pricing of Local Coffee Shop to Increase Consumer Satisfaction and Loyalty," *J. Tek. Pertan. Lampung (Journal Agric. Eng.*, vol. 13, no. 2, p. 592, 2024, doi: 10.23960/jtep-l.v13i2.592-605.

[10] S. M. M. MangusStephanie M, "Personal selling and sales management abstracts," *J. Pers. Sell. Sales Manag.*, vol. 44, no. 4, pp. 343–354, 2024, doi: 10.1080/08853134.2024.2415453.

[11] M. A. Majid, P. D. Saputri, and S. Soehardjoepri, "Stock Market Index Prediction using Bi-directional Long Short-Term Memory," *J. Appl. Informatics Comput.*, vol. 8, no. 1, pp. 55–61, 2024, doi: 10.30871/jaic.v8i1.7195.

[12] Y. Chen, "Research on Machine Learning-based Prediction of Coffee Futures Prices," *Highlights Sci. Eng. Technol.*, vol. 92, pp. 199–209, 2024, doi: 10.54097/8c9t9n30.

[13] M. Miranda and S. Sriani, "Implementation of K-Means Clustering in Grouping Sales Data at Zura Mart," *J. Appl. Informatics Comput.*, vol. 9, no. 2, pp. 547–555, 2025, doi: 10.30871/jaic.v9i2.9160.

[14] R. Shams, S. Chatterjee, and R. Chaudhuri, "Developing brand identity and sales strategy in the digital era: Moderating role of consumer belief in brand," *J. Bus. Res.*, vol. 179, no. April, p. 114689, 2024, doi: 10.1016/j.jbusres.2024.114689.

[15] V. R. Danestiara, M. Marwondo, and N. N. Azkiya, "Prediction of Inhibitor Binding Affinity and Molecular Interactions in Mpro Dengue Using Machine Learning," *JITK (Jurnal Ilmu Pengetah. dan Teknol. Komputer)*, vol. 10, no. 3, pp. 461–468, 2025, doi: 10.33480/jitk.v10i3.5994.

[16] S. Febriani, V. Wati, Y. Wijayanti, and I. Siswanto, "Twitter Sentiment Analysis on Digital Payment in Indonesia Using Artificial Neural Network," *J. Appl. Informatics Comput.*, vol. 9, no. 2, pp. 526–533, 2025, doi: 10.30871/jaic.v9i2.8988.

[17] F. Mercaldo, L. Brunese, F. Martinelli, A. Santone, and M. Cesarelli, "Experimenting with Extreme Learning Machine for Biomedical Image Classification," *Appl. Sci.*, vol. 13, no. 14, pp. 1–20, 2023, doi: 10.3390/app13148558.

[18] M. Vinoth Kumar, G. Kiran Kumar, A. M. Ali, S. K. Rajesh Kanna, and K. Muthu Lakshmi, "Analysis of Extreme Learning Machine Based on Multiple Hidden Layers," *Int. J. Intell. Syst. Appl. Eng.*, vol. 12, no. 95, pp. 96–103, 2024, [Online]. Available: https://www.ijisae.org/index.php/IJISAE/article/view/4208

[19] A. P. Ariyanti, M. I. Mazdadi, A.- Farmadi, M. Muliadi, and R. Herteno, "Application of Extreme Learning Machine Method With Particle Swarm Optimization to Classify of Heart Disease," *IJCCS (Indonesian J. Comput. Cybern. Syst.*, vol. 17, no. 3, p. 281, 2023, doi: 10.22146/ijccs.86291.

[20] R. Fredyan, M. R. N. Majiid, and G. P. Kusuma, "Spatiotemporal Analysis for Rainfall Prediction Using Extreme Learning Machine Cluster," *Int. J. Adv. Sci. Eng. Inf. Technol.*, vol. 13, no. 6, pp. 2240–2248, 2023, doi: 10.18517/ijaseit.13.6.18214.

[21] C. D. Nariyana, M. Idhom, and Trimono, "Prediction of Purchase Volume Coffee Shops in Surabaya Using Catboost with Leave-One-Out Cross Validation," *J. Ilm. Tek. Elektro Komput. dan Inform.*, vol. 11, no. 1, pp. 124–138, 2025, doi: 10.26555/jiteki.v11i1.30610.

[22] W. K. Setiadi, V. R. Prasetyo, and F. D. Kartikasari, "Comparison of Extreme Learning Machine Methods and Support Vector Regression for Predicting Bank Share Prices in Indonesia," *Teknika*, vol. 13, no. 2, pp. 219–225, 2024, doi: 10.34148/teknika.v13i2.856.

[23] D. Das and S. Chakrabarti, "An Extreme Learning Machine Approach for Forecasting the Wholesale Price Index of Food Products in India," *Pertanika J. Sci. Technol.*, vol. 31, no. 6, pp. 3179–3198, 2023, doi: 10.47836/pjst.31.6.30.

[24] J. M. A. C. Permata and M. Habibi, "Autoregressive Integrated Moving Average (ARIMA) Models For Forecasting Sales Of Jeans Products," *J. Inform. dan Teknol. Inf.*, vol. 20, no. 1, pp. 31–40, 2023, doi: 10.31515/telematika.v20i1.7868.

[25] M. Khairunisa, D. M. S. A. Putri, and I. G. N. L. Wijayakusuma, "Comparison of Machine Learning Methods for Menstrual Cycle Analysis and Prediction," *J. Appl. Informatics Comput.*, vol. 9, no. 2, pp. 348–353, 2025, doi: 10.30871/jaic.v9i2.9076.

[26] S. Aisyah, N. Ulinnuha, and A. Hamid, "Penerapan Extreme Learning Machine Dalam Meramalkan Harga Minyak Sawit Mentah," *KUBIK J. Publ. Ilm. Mat.*, vol. 7, no. 2, pp. 97–105, 2023, doi: 10.15575/kubik.v7i2.20460.

[27] S. Sulandri, A. Basuki, and F. A. Bachtiar, "Metode Deteksi Intrusi Menggunakan Algoritme Extreme Learning Machine dengan Correlation-based Feature Selection," *J. Teknol. Inf. dan Ilmu Komput.*, vol. 8, no. 1, p. 103, 2021, doi: 10.25126/jtiik.0813358.

[28] C. Wiedyaningsih, E. Yuniarti, and A. Fadilla, "Forecasting Cardiovascular Drug Demand Using Triple Exponential Smoothing Additive," *Indones. J. Glob. Heal. Res.*, vol. 7, no. 1, pp. 29–34, 2025, doi: 10.37287/ijghr.v2i4.250.