Long Short-Term Memory Method Based on Normalization Data For Forecasting Analysis of Madura Ginger Selling Price

Devie Rosa Anamisa[1], Fifin Ayu Mufarroha[1], Achmad Jauhari[1], Muhammad Yusuf [1], Bain Khusnul Khotimah[1] and Ahmad Farisul Haq[1]

[1]Informatics Engineering Department, University of Trunojoyo Madura, Bangkalan, Indonesia
devros_gress@trunojoyo.ac.id

Abstract. Forecasting is a method for estimating a future value using past data. The selling price of Madura ginger needs a forecasting analysis to predict future prices because, until now, the selling price has increased significantly. This analysis aims to increase trade business competition and maintain sales objectives related to financing, revenue planning, and marketing. In this study, the forecasting analysis system uses the Long Short-Term Memory (LSTM) method. LSTM is one of the forecasting methods with the development of a neural network that can be used for modeling time series data collected according to a time sequence within a specific time. This research contributes to forecasting ginger's selling price in Madura using LSTM with improved model performance using max-min normalization for the preprocessing process. Max-min normalization eliminates data redundancy by converting a data set to a scale from 0 (min) to 1 (max) to make the data consistent. And for this study's forecasting analysis, use error parameters including Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE). Based on the simulation results of ginger spice sales price data in Madura, it was obtained that the 2019 ginger selling price prediction was 250 data with a value of RMSE 1431.71 and MAPE 9.57. This shows that the results of the LSTM modeling have shown excellent performance in predicting training and testing the selling price of ginger so that the prediction of the selling price of ginger in 2020 can increased tolerance for time-series data and the accuracy with normalization model.

Keywords: Forecasting Analysis, Ginger Selling Price, Long Short-Term Memory, Max-Min Normalization.

1 Introduction

Ginger is one of the spice plants that often experience fluctuations in selling prices at the farmer and consumer levels. This price fluctuation is considered one of the causes of inflation in Indonesia [1]. This ups and downs of ginger prices greatly affect the profits of ginger spice farmers in re-producing. When prices are low, farmers cannot produce optimally, so the impact that consumers will feel is an increase in prices be

M. Hartono et al. (eds.), *Proceedings of the 4th International Conference on Informatics, Technology and Engineering 2023 (InCITE 2023)*, Atlantis Highlights in Engineering 21,
https://doi.org/10.2991/978-94-6463-288-0_52

abundant, it is not profitable for ginger farmers because of the nature of the ginger plant which is not durable [2] so it cannot be stored for a long time. Therefore, it is necessary to have a forecasting analysis system to estimate sales accurately from time to time so that a production plan can be made according to the estimated selling price. Forecasting analysis is an analytical process to predict future conditions by testing past data [3].

In several previous studies, there are several approaches to forecasting methods. That have been used for analysis, such as research conducted by [4] in 2011 regarding shanghai stock forecasting with data from June 2006 to November 2009, where the Backpropagation method influenced by factors such as economics, and politics, produces experimental effects with short-term prediction data by Backpropagation neural networks is valid. Then the development of research by [5] in 2019 regarding the application of backpropagation neural networks for forecasting rice sales and the results obtained from the Means Square Error (MSE) value using the linear regression forecasting method value of 0.214 while using ANN it produces 0.00099713. MSE in this study is used to compare the value that comes out in the output layer with the actual or target value. Meanwhile, in 2020, research by [6] by using a different method, namely Fuzzy C-Mean based on the ANFIS method for wind speed estimates, and this results in 720 data the RMSE value is 0.7. This shows that the RMSE needs to be reduced to a very low value from the actual value so that ANFIS and Fuzzy C-Mean are able to group each cluster with a range of values to produce better forecasts. Therefore, it was developed by research conducted by [7] in 2022 regarding weather forecasts using the support vector machine (SVM) method in determining the exact value of weather parameters and further in the future. However, this study shows that the process that has been applied still indicates a relatively high error rate so that it is still not able to improve knowledge or performance for time series data [8]. Meanwhile, in the same year, research by [9] using LSTM for electrical power load forecasting, where the LSTM forecast model is used to get more accurate power load prediction results. The data set is divided into three parts, including 86.5% training data, 3.5% verification data and 10% test data. From these data, the results can be predicted using the LSTM prediction model. The results of single point load forecasting produce MAPE 1.806. In addition, in research by [10] regarding next word prediction using the LSTM method and using 180 Indonesian destinations from nine provinces for training data. From this prediction, the LSTM method has shown an accuracy of 75% with 200 epochs and an MAE value of 0.068 and an RMSE of 0.99.

Therefore, this study tries to apply the LSTM method approach to the analysis of ginger selling price forecasting in Madura. This is because the LSTM method is one of the forecasting methods that has a high level of accuracy for predicting data in the form of time series so that it is able to overcome the vanishing gradient or the state of the gradient value is 0 or close to 0 with the gate mechanism[11]. The purpose of this study is to conduct a forecasting analysis for the selling price of spiced ginger by applying the LSTM method to estimate the model with epoch variations to produce an estimate of the selling price of ginger with accurate results with small errors by taking into account the RMSE level error and Mean Absolute Percentage Error (MAPE) and can be used as a basis for the development of further research.

2. Research Method

According to the research flow shown in Figure 1, this research is generally divided into three stages: the data preprocessing stage, the LSTM model creation stage, and the simulation results evaluation stage. The data used in this forecasting session is data on the selling price of ginger spices from 2015-2019, as many as 250 datasets from various regions of ginger-producing farmers in Madura that have been collected at the Madura Agriculture Service. The first process is data collection. The dataset can be seen in Table 1. After collecting the data, the second step is to perform preprocessing by normalizing the data. This study uses the Max-Min Normalization method. Data normalization is the process of scaling the attribute values of the data so that they fall within a specific range [12]. Normalization calculation process to produce output with a data range between 0 and 1 [13]. The data normalization process is shown in Equation (1).

$$X' = \frac{(X - X_{min})}{X_{max} - X_{min}} \tag{1}$$

where $X'$is the normalized data, X is the original data, X_{max} is the maximum value of the original data, and X_{min} is the minimum value of the original data. The results of the normalization process can be seen in Table 2. In Table 2, there are “Min” is the smallest value of the dataset. The purpose of normalization is to increase accuracy in forecasting[14]. With Max-Min normalization in the preprocessing process, it can group data attributes that have good quality, preventing data inconsistencies or missing values. So, Min-Max Normalization can maintain the relationship between the original data values. And this method will not encounter an "out of bounds" error if the new data input case is outside the actual data range. However, there are obstacles when in an area in Madura, the ginger selling price data is not recorded correctly, so the characteristics of the data cannot be identified in the pattern. It is necessary to re-record it in a particular month by inputting the number zero to determine the forecasting performance and the level of achievement of adequate accuracy. Then the denormalization process to return data values to their original or actual values will be carried out based on the forecasting results. The method of calculating the denormalization of the data is shown in Equation (2). The denormalization process was carried out on 74% of the 250 datasets, about 184 data and 26% of which are test data of 26 data. And the results of the denormalization can be seen in Table 3 with 184 training data.

$$X = \left(X'(X_{max} - X_{min})\right) + X_{min} \tag{2}$$

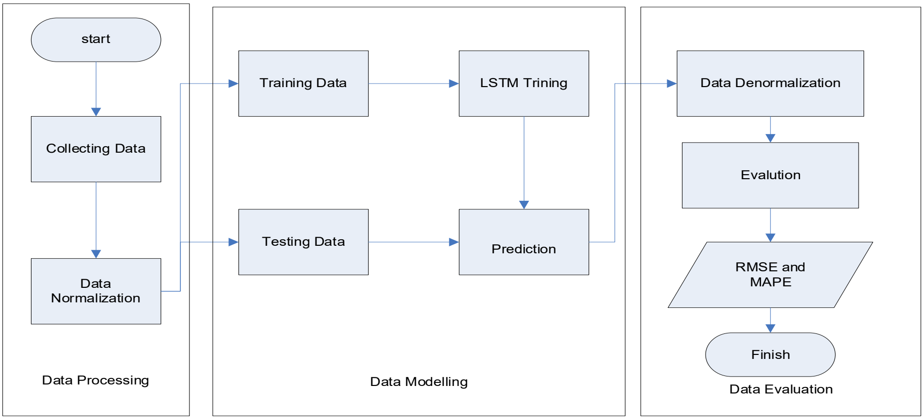

Fig. 1. Research methodology

Table 1. Dataset of ginger selling price

No	Region	Year	Selling Price of Ginger	No	Region	Year	Selling Price of Ginger
1	Arosbaya	2015	11000	126	Karang Penang	2019	14700
2	Bangkalan	2016	13100	127	Kedundung	2015	15200
3	Blega	2017	14900	128	Rubaru	2016	11500
4	Burneh	2018	10900	129	Sapeken	2017	14600
5	Galis	2019	14500	130	Saronggi	2018	14100
6	Ambunten	2015	15200	131	Sumenep	2019	13900
…	….	…	…	….	…	…	…
…	….	…	…	….	…	…	…
119	Arjasa	2017	12800	244	Paragaan	2018	12114
120	Batang-batang	2018	11400	245	Pasongsongan	2019	13764
121	Batuan	2019	10700	246	Omben	2015	11456
122	Batuputih	2015	13300	247	Pangarengan	2016	13323
123	Banyuates	2016	13000	248	Robatal	2017	13787
124	Camplong	2017	15400	249	Sampang	2018	10553
125	Jrengik	2018	12800	250	Sokobanah	2019	13672

The next step after carrying out the denormalization process is the training process. However, before starting the training, this study uses the LSTM model for the ginger selling price forecasting process. The benefit of LSTM for forecasting the selling price of ginger is that it can search for accurate solutions and better learning abilities for large data sets, and can improve forecasting accuracy. Therefore, making LSTM modeling with one hidden layer and one output, determining the number of batch size is 32, the optimizer using epoch value from 10 to 40, and choosing the loss function using Mean Squared Error (MSE) with a value of 0.001. The LSTM model is one model that has a variant of the Recurrent Neural Network (RNN) unit [15]. The LSTM method generally consists of a cell, input gate, output gate, and forget gate[16]. LSTM is particularly well suited for classifying, processing, and making predictions based on time series data as there may be a rarity of unknown duration among significant events in time series. The general architecture of the LSTM consists of a memory cell, input gate, output gate, and forget gate, which can be seen in Figure 2 [16][17]. The incoming data on the forget gates will be processed according to the information, and the selected data will be stored in the memory cell. Equation (3) describes the working principle. At the same time, the input gates have two gates that use the sigmoid activation function to update information and the tanh activation function to store the new value in the memory cell. This can be illustrated in equations (4) and (5). Equation (6) is the result of the combined values at the input gate. Forget gates will replace the memory cell value with cell gates. At the output gates, there are two gates to decide the value to be issued with the sigmoid activation function and store the matter using the tanh activation function. This is formulated in equations (7) and (8) [18][19]. The training process uses training data. Weight and bias will be updated for a suitable model [20]. After the training process, the next step is the validation process to evaluate the two models from the training results.

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_t) \tag{3}$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{4}$$

$$\hat{c}_t = tanh(W_c[h_{t-1}, x_t] + b_c) \tag{5}$$

$$c_t = f_t * c_{t-1} + i_t * \hat{c}_t \tag{6}$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \tag{7}$$

$$h_t = o_t \tanh(c_t) \tag{8}$$

where f_t is calculation results on the forget gate. The forget gate in the LSTM unit functions as a determinant of status information on cells to be discarded or not from the model. W_f is weight matrix on forget gate (U_f). i_t is calculation results at the input gate are used to determine the cell state that meets the requirements for updating where the value is selected through the sigmoid layer. W_i is weight matrix at the input gate (U_i). C_t is the calculation result for the new value added in the cell state, which functions to update information on the cell state (cell state). W_c is the weight matrix in the cell state or called U_c. Furthermore, to set the number of cell states at the current time that must be discarded is O_t with W_o is the weight matrix at the Output gate (U_o). While the final output value is processed using the calculation by the tanh function by the cell state multiplied by the gate output result from the sigmoid layer (h_t).

Table 2. Result of normalization data

No	Selling Price of Ginger	No	Selling Price of Ginger	No	Selling Price of Ginger	No	Selling Price of Ginger	No	Selling Price of Ginger
Min	10500								
Min	15400								
1	0,10204	51	0	101	0,2857	151	0,1020	201	0,4826
2	0,53061	52	0,87755	102	0,1224	152	0,9795	202	0,3195
3	0,89795	53	0,95918	103	0,6122	153	0,1632	203	0,5936
4	0,08163	54	0,48979	104	0,3673	154	0,0408	204	0,3283
5	0,81632	55	0,57142	105	0,4897	155	0,5918	205	0,5553

6	0,95918	56	0,24489	106	0,9387	156	0,8367	206	0,9032
7	0,02040	57	0,69387	107	0,5306	157	0,7346	207	0,5773
8	0,91836	58	0,04081	108	0,6938	158	0,1836	208	0,1128
9	0,73469	59	0,10204	109	0,8367	159	0,5102	209	0,0130
10	0,38775	60	0,44898	110	0,9387	160	0,0204	210	0,8361
…	….	…	…	…		…	…	…	…
40	0,2244	90	0,10204	140	0.8610	190	0,1785	240	0,6689
41	0,5918	91	0,95918	141	0.0650	191	0,7348	241	0,4320
42	0,4489	92	0,30612	142	0.7912	192	0,3693	242	0,3216
43	0	93	0,40816	143	0.1414	193	0,2783	243	0,2148
44	0,9795	94	0,91836	144	0.6779	194	0,8559	244	0,3293
45	0	95	0,32653	145	0.4808	195	0,4991	245	0,6661
46	0,1020	96	0,24489	146	0.9282	196	0,5148	246	0,1951
47	0,0612	97	0,28571	147	0.5317	197	0,6132	247	0,5761
48	0,7346	98	0,59183	148	0.6281	198	0,5836	248	0,6708
49	0,1428	99	0,30612	149	0.4111	199	0,2932	249	0,0108
50	0,2653	100	0,06122	150	0.006	200	0,8179	250	0,6473

Then the last step is the testing process for performance measurement of the model. Evaluation of this model is used to determine the accuracy of the model made. In this study, the measurement of model performance is based on MAPE calculations. MAPE is the percentage of error resulting from the average value of the difference between the actual data and the forecasted data [21][22]. The smaller the MAPE value of a model, it can be said that the forecasting model made the better the performance. With the MAPE formula in equation (9). Besides MAPE, this study also measures performance with RMSE. RMSE is a measurement method by measuring the difference in the value of the prediction of a model as an estimate of the observed value [23][24]. The accuracy of the measurement error estimation method is indicated by the presence of a small RMSE value. The estimation method that has a smaller RMSE is said to be more accurate than the estimation method that has a larger RMSE. The way to calculate RMSE is to subtract the actual value from the forecast value then square and add up the overall results and then divide by the numbers of data. The results of these calculations are then recalculated to find the value of the square root, which can be seen in equation (10)[25].

Table 3. Result of Denormalization Data

No	Selling Price of Ginger	No	Selling Price of Ginger	No	Selling Price of Ginger	No	Selling Price of Ginger	No	Selling Price of Ginger
Min	10500								
Min	15400								
1	13060.017	51	10629.5591	101	12560.68	151	12507.28	201	12865.00
2	12956.074	52	11802.8908	102	12280.36	152	12809.50	202	12066.00
3	12729.084	53	12973.3437	103	12834.76	153	12893.05	203	13409.00
4	12912.777	54	12987.1512	104	12795.37	154	12620.55	204	12109.00
5	13376.331	55	13124.9902	105	12645.60	155	12676.19	205	13221.00

6	13521.699	56	13001.7108	106	12795.31	156	12743.36	206	14926.00
7	13468.291	57	12985.7359	107	12970.10	157	12451.88	207	13329.00
8	13503.657	58	13115.7619	108	12643.53	158	12607.02	208	11053.00
9	13161.669	59	12954.3955	109	12849.81	159	11810.02	209	10564.00
10	13016.823	60	12819.7911	110	12896.08	160	11571.54	210	14597.00
…	….	…	…	…		…	…	…	…
40	12478.228	90	12420.7467	140	12573.72	190	11375.00	240	13778.00
41	12578.569	91	12287.0137	141	12493.08	191	14101.00	241	12617.00
42	12136.682	92	12071.0298	142	12675.93	192	12310.00	242	12076.00
43	12171.170	93	12479.7938	143	12666.92	193	11864.00	243	11553.00
44	12235.226	94	12610.8357	144	13054.42	194	14694.00	244	12114.00
45	12255.727	95	12504.4147	145	12715.12	195	12946.00	245	13764.00
46	11585.763	96	12927.7624	146	12399.13	196	13023.00	246	11456.00
47	12135.354	97	12517.6311	147	11886.99	197	13505.00	247	13323.00
48	12306.932	98	12729.0291	148	11322.43	198	13360.00	248	13787.00
49	12185.317	99	12841.2694	149	11943.26	199	11937.00	249	10553.00
50	11700.662	100	12801.3829	150	12295.39	200	14508.00	250	13672.00

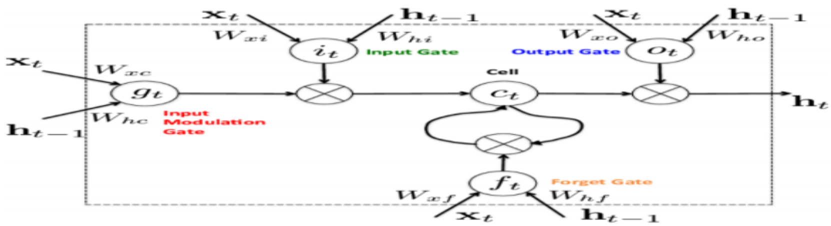

Fig. 2. Architecture of LSTM General

$$MAPE = 100\% * \sum_{t=1}^{n} \left|\frac{A_t - F_t}{A_t}\right| \tag{9}$$

$$RMSE = \sqrt{\sum_{t=1}^{n} \frac{(A_t - F_t)^2}{n}} \tag{10}$$

where n is the amount of data, At is the actual value, and F_t is the forecast value. MAPE and RMSE are techniques that are easy to implement and have been widely used in various studies related to prediction or forecasting.

3. Result and Discussion

This research was conducted to obtain the results of forecasting the selling price of ginger in Madura in 2020 using the ginger selling price dataset from 2015 to 2019, as shown in Table 1. Previous and new inputs pass them through the sigmoid layer or the input gate (i_t). This gate returns a value of 0 or 1. Then the value of the input gate is multiplied by the output of the candidate layer (C_t) and calculates f_t and controls how many states pass to the production and works in the same way as other gates or is called o_t. And finally generate a new cell state (h_t). The final results of calculations using LSTM as an example in one month can be seen in Table 4.

Table 4. Results of the training process with LSTM

$h_{t-1} = 0$	Week								
$c_{t-1} = 0$	I	II	III	IV	f_t	i_t	$\hat{C}_t$	o_t	h_t
Data I	1	0.76	0.55	1	0.93	0.97	0.98	0.91	0.68
U_f	0.65	0.35	0.8	0.7					
Data II	0.15	1	0.10	0.15	0.72	0.90	0.87	0.67	0.61
U_i	0.34	0.63	0.74	0.95					
Data III	0.56	0	1	0.56	0.88	0.95	0.88	0.83	0.81
U_c	0.85	0.77	0.23	0.45					
Data IV	0.84	0.75	0	0.84	0.88	0.97	0.97	0.88	0.88
U_o	0.95	0.13	0.25	0.6					
Data V	0	0.47	0.97	0	0.76	0.91	0.73	0.65	0.65

Experiment with 40 iterations or epochs for the ginger selling price dataset, and the output is shown in Table 5. Then experimented with the correlation of the output results in accordance with several predetermined epochs. shown in table 5 with 100 nodes of LSTM or neuron units. Correlation is the relationship between variables, which is analyzed to see how strong the relationship between variables is and in what direction. The direction of the relationship is indicated by a positive or negative relationship. The best forecasting is based on the level of prediction error, where the smaller the resulting error rate, the more precise a method is in predicting. Therefore, in this study to determine the number of neurons, an experiment was carried out to obtain optimal results in predictions that can be seen from the error value, or what is called an epoch. Epoch is a step taken in the neural network learning process, which affects the amount of the learning process and stops at the predetermined epoch value. In this study the epochs used for the experiment were 10, 15, 20, 25, 30, 35, and 40. To find out the exact epoch value, the smallest loss value can be seen, where the loss value is the RMSE value and Adam's optimization is used. The number of epochs that are too few results in the network being formed to be too general, meaning the network's capabilities in recognizing patterns too little or not at all. While too many epochs will cause the network to experience an overfit condition (the network is too specific to the training data), it can be seen in table 6 that the best results are not at the largest epoch value. From Table 6 it can be seen that the smallest RMSE value is found in the data combination of epoch 15 and the smallest MAPE is at epoch 40. The largest RMSE value is found in the combination of data from epoch 40 and the largest MAPE is at epoch 10. While Figure 3 and Figure 4 shows a gap between predictions and expectations. generated from the LSTM model for a certain epoch produces the smallest and largest RSME values, where the predicted value is the result of the process calculated by LSTM. And Figure 5 shows the prediction and expectation results at epoch 100 and the window size is 20 with the expected value obtained from the dataset.

Table 5. Result of predicted values and experimental actual values for 2019

Month	Prediction	Expectation	Date	Prediction	Expectation
January	12563.9	15300	July	12520.9	10842
	12723.8	15300		12251.1	12400

	13212.8	10800		12469.0	10803
	12660.3	13200		12531.7	14100
February	12580.2	13672	August	13034.6	15000
	12428.5	10600		12841.7	11900
	12613.0	15300		13287.3	13207
	13503.2	14101		13449.4	13300
March	12950.7	14600	September	13313.9	13900
	13164.2	14400		13138.2	14500
	12273.9	10500		13068.2	12100
	12262.6	11800		12639.3	11400
April	11695.8	10709	October	11926.6	12000
	11909.2	11100		11806.3	12700
	12015.9	11894		12096.2	11375
	12388.9	13296		11721.3	10800
May	12609.8	10600	November	12426.1	13787
	12367.9	13900		12759.9	15400
	12299.5	14915		12696.8	11400
	12552.1	12800		12599.3	14400
June	12336.2	12000	December	12723.4	11000
	12575.7	14300		12453.3	13000
	12251.6	10564		12313.1	12728
	12208.5	10612		12599.5	13000

Table 6. Prediction correlation results of selling price of ginger with MAPE and RMSE

No	Epoch	Error Testing	
		MAPE	RMSE
1	10	9.76	1447.24
2	15	9.64	1431.71
3	20	9.74	1444.97
4	25	9.63	1435.79
5	30	9.60	1432.15
6	35	9.66	1438.39
7	40	9.57	1451.15

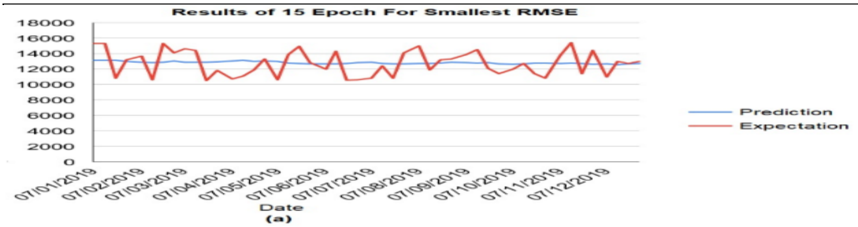

Fig. 3. Smallest RMSE

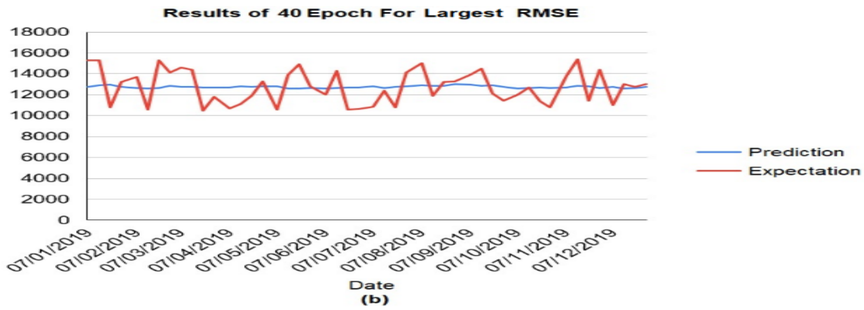

Fig. 4. Largest RMSE

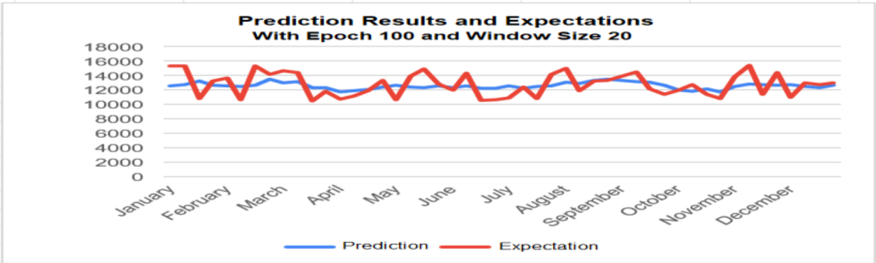

Fig. 5. Graph of result of 2019 data prediction and expectation with epoch 100

4. Conclusion

From the results of this study, a prediction model with the LSTM method has been produced which can be used quite well to predict the selling price of ginger for the following year. Testing the model with several experiments to find the best model of the epoch with the smallest RMSE and MAPE values so that it can be concluded that there is no large gap between expectations and predictions, which means that the level of accuracy is quite high. Correlation test with 250 ginger selling price data for 2019 and it was concluded that the predictive variables and RMSE and MAPE values have a significant correlation with each other with RMSE values of 1431.71 and MAPE 9.57 which are the smallest values so it is known that the model built is quite good to predict the selling price of ginger and the prediction results can be used as a basis for decision making. Suggestions that can be given for further research are optimizing the LSTM architecture or others and then comparing it.

Acknowledgment

Thank you to the University of Trunojoyo Madura, for the opportunity to do research with contract number 158 /UN46.4.1/PT.01.03/2022. And the author also thanks the Department of Agriculture and Horticulture in the Sumenep office for their assistance in providing data in this research

References

1. Kaharuddin, Kusrini, Vera Wati, Elvis Pawan, Patmawati Hasan. "Classification of Spice Types Using K-Nearest Neighbor Algorithm". International Conference on Information and Communications Technology (ICOIACT), 285–290 (2019), doi: 10.1109/ICOIACT46704.2019.8938515.
2. L. W. Nagel and C. C. McAndrew, "Why SPICE is just as good and just as bad for IC design as it was 40 years ago," Eur. Solid-State Device Res. Conf., vol. 2018-Septe, 170–173 (2018), doi: 10.1109/ESSDERC.2018.8486875.
3. C. Zheng and J. Zhu, "Research on stock price forecast based on gray relational analysis and ARMAX model," 2017 IEEE Int. Conf. Grey Syst. Intell. Serv. GSIS 2017, 145–148 (2017), doi: 10.1109/GSIS.2017.8077689.
4. Y. Ma, C. Yu, and C. Xia, "Applied research on stock forcasting model based on BP neural network," Proc. 2011 Int. Conf. Electron. Mech. Eng. Inf. Technol. EMEIT 2011, vol. 9, 4578–4580(2011), doi: 10.1109/EMEIT.2011.6023120.
5. M. Aritonang and D. J. C. Sihombing, "An Application of Backpropagation Neural Network for Sales Forecasting Rice Miling Unit," 2019 Int. Conf. Comput. Sci. Inf. Technol. ICoSNIKOM 2019, 7–10 (2019), doi: 10.1109/ICoSNIKOM48755.2019.9111612.
6. M. Ramesh Babu, A. Q. H. Badar, and S. Balasubramani, "Fuzzy-C means clustering based ANFIS wind speed forecast," 2020 21st Natl. Power Syst. Conf. NPSC 2020, (2020), doi: 10.1109/NPSC49263.2020.9331828.
7. Q. Zhao, Y. Liu, W. Yao, and Y. Yao, "Hourly Rainfall Forecast Model Using Supervised Learning Algorithm," IEEE Trans. Geosci. Remote Sens., vol. 60, 1–9 (2022), doi: 10.1109/TGRS.2021.3054582.
8. J. F. L. De Oliveira, E. G. Silva, and P. S. G. De Mattos Neto, "A Hybrid System Based on Dynamic Selection for Time Series Forecasting," IEEE Trans. Neural Networks Learn. Syst., vol. 33, no. 8, 3251–3263(2022), doi: 10.1109/TNNLS.2021.3051384.
9. C. Cui, M. He, F. Di, Y. Lu, Y. Dai, and F. Lv, "Research on Power Load Forecasting Method Based on LSTM Model," Proc. 2020 IEEE 5th Inf. Technol. Mechatronics Eng. Conf. ITOEC 2020, no. Itoec, 1657–1660(2020), doi: 10.1109/ITOEC49072.2020.9141684.
10. S. Agarwal, Sukritin, A. Sharma, and A. Mishra, "Next Word Prediction Using Hindi Language," Lect. Notes Networks Syst., 356(1), 99–108(2022), doi: 10.1007/978-981-16-7952-0_10.
11. S. Kumar and D. Ningombam, "Short-Term Forecasting of Stock Prices Using Long Short Term Memory," Proc. - 2018 Int. Conf. Inf. Technol. ICIT 2018, 182–186(2018), doi: 10.1109/ICIT.2018.00046.
12. B. S. Kwon, R. J. Park, S. W. Jo, and K. Bin Song, "Analysis of short-term load forecasting using artificial neural network algorithm according to normalization and selection of input data on weekdays," Asia-Pacific Power Energy Eng. Conf. APPEEC, vol. 2018-Octob, 280–283(2018), doi: 10.1109/APPEEC.2018.8566293.
13. V. Gajera, Shubham, R. Gupta, and P. K. Jana, "An effective Multi-Objective task scheduling algorithm using Min-Max normalization in cloud computing," Proc. 2016 2nd Int. Conf. Appl. Theor. Comput. Commun. Technol. iCATccT 2016, 812–816(2017), doi: 10.1109/ICATCCT.2016.7912111.
14. C. Yu, Z. Li, Z. Yang, X. Chen, and M. Su, "A feedforward neural network based on normalization and error correction for predicting water resources carrying capacity of a city," Ecol. Indic., 118(July), 106724(2020), doi: 10.1016/j.ecolind.2020.106724.

15. J. Qin et al., "Multi-task short-term reactive and active load forecasting method based on attention-LSTM model," Int. J. Electr. Power Energy Syst., 135(May 2020), 107517(2022), doi: 10.1016/j.ijepes.2021.107517.
16. D. Fan, H. Sun, J. Yao, K. Zhang, X. Yan, and Z. Sun, "Well production forecasting based on ARIMA-LSTM model considering manual operations," Energy, vol. 220, 119708(2021), doi: 10.1016/j.energy.2020.119708.
17. Y. S. Chang, H. T. Chiao, S. Abimannan, Y. P. Huang, Y. T. Tsai, and K. M. Lin, "An LSTM-based aggregated model for air pollution forecasting," Atmos. Pollut. Res., 11(8), 1451–1463(2020), doi: 10.1016/j.apr.2020.05.015.
18. S. Urolagin, N. Sharma, and T. K. Datta, "A combined architecture of multivariate LSTM with Mahalanobis and Z-Score transformations for oil price forecasting," Energy, vol. 231, 120963(2021), doi: 10.1016/j.energy.2021.120963.
19. K. He, L. Ji, C. W. D. Wu, and K. F. G. Tso, "Using SARIMA–CNN–LSTM approach to forecast daily tourism demand," J. Hosp. Tour. Manag., 49(July), 25–33(2021), doi: 10.1016/j.jhtm.2021.08.022.
20. G. Accarino et al., "A multi-model architecture based on Long Short-Term Memory neural networks for multi-step sea level forecasting," Futur. Gener. Comput. Syst., vol. 124, 1–9(2021), doi: 10.1016/j.future.2021.05.008
21. V. K. Sudarshan, M. Brabrand, T. M. Range, and U. K. Wiil, "Performance evaluation of Emergency Department patient arrivals forecasting models by including meteorological and calendar information: A comparative study," Comput. Biol. Med., 135(June), 104541(2021), doi: 10.1016/j.compbiomed.2021.104541.
22. M. L. Shen, C. F. Lee, H. H. Liu, P. Y. Chang, and C. H. Yang, "Effective multinational trade forecasting using LSTM recurrent neural network," Expert Syst. Appl., 182(May), 115199(2021), doi: 10.1016/j.eswa.2021.115199.
23. D. C. Wu, B. Bahrami Asl, A. Razban, and J. Chen, "Air compressor load forecasting using artificial neural network," Expert Syst. Appl., 168(November), 114209(2021), doi: 10.1016/j.eswa.2020.114209.
24. E. Dave, A. Leonardo, M. Jeanice, and N. Hanafiah, "Forecasting Indonesia Exports using a Hybrid Model ARIMA-LSTM," Procedia Comput. Sci., 179(2020), 480–487 (2021), doi: 10.1016/j.procs.2021.01.031.
25. H. D. Nguyen, K. P. Tran, S. Thomassey, and M. Hamad, "Forecasting and Anomaly Detection approaches using LSTM and LSTM Autoencoder techniques with the applications in supply chain management," Int. J. Inf. Manage., 57(November), 102282(2021), doi: 10.1016/j.ijinfomgt.2020.102282.