ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

PALM OIL PRODUCTION PREDICTION USING LONG SHORT-TERM MEMORY METHOD

[1]HANDRIZAL, [2]DEWI SARTIKA GINTING, [3]HERRIYANCE, [4]SYABRINA RAMADHANI KAMAL

[1,2,3]Department of Computer Science, Faculty of Computer Science and Information Technology, Universitas Sumatera Utara, Jl. University No. 9-A, Medan 20155, Indonesia

E-mail: handrizal@usu.ac.id

ABSTRACT

Palm oil production is one of the important indicators in the agribusiness sector. Accurate prediction of palm oil production helps better planning and decision-making. This study's problem is building a palm oil production prediction model using Long Short-Term Memory (LSTM). The LSTM model consists of four layers with 50 memory units and 20% dropout. The model is compiled using the RMSprop optimizer and the RMSE loss function. The evaluation results show that the LSTM model provides palm oil production predictions with RMSE of 0.1360 for total FFB received, 0.1279 for total FFB processed, and 0.1279 for total CPO which shows good potential. The contribution of this research can provide knowledge about predicting palm oil production so that companies can plan more efficient production.

Keywords: *Long Short-Term Memory (LSTM), Palm Oil Production Prediction, MinMaxScaler, Root Mean Square Error (RMsE), Deep Learning*

1. INTRODUCTION

Palm oil production is one of the main economic pillars for producing countries such as Indonesia and Malaysia. Palm oil production has increased significantly in recent decades to meet global demand. However, commodity price fluctuations, climate change, and other factors often affect production, posing challenges for industry players in planning production and distribution efficiently.

PT X, located in Riau, processes palm fruit bunches into crude palm oil. The company faces fluctuation in the number of incoming Palm Fruit Bunches, causing uncertainty in Crude Palm Oil production. This has an impact on machine maintenance time, oil stocks, and the company's cash flow.

One of the methods to predict Crude Palm Oil production is data mining techniques. Data mining extracts information from large data sets to understand existing patterns and characteristics, helping companies make decisions [1]. The Long Short-Term Memory (LSTM) algorithm, a type of artificial neural network, is often used in time series data processing because of its ability to remember important information over a long period [2].

In some studies, Long Short-Term Memory is used to predict production quantities, as was done by Anwar [3], this paper surveys the short-term road traffic forecast algorithms based on the long-term term memory (LSTM) model of deep learning. The algorithms developed in the last three years are studied and analyzed. This provides a thorough description of the algorithms rather than their marginal description as performed in the existing surveys that focus on general deep learning algorithms. The chosen algorithms are classified depending upon the use of LSTM in combination with other techniques for processing input data features toward a final traffic forecast. However, this study only surveys short-term road traffic forecasting algorithms based on short-term memory (LSTM) deep learning models without any RMSE evaluation.

Research by Yuwen [4] proposed a greenhouse climate prediction model. The focus of this research is on six climate factors affecting plant growth, including temperature, humidity, lighting, carbon dioxide concentration, soil temperature, and soil moisture, and promotes the GCP_lstm model for greenhouse climate prediction. However, this study only focuses on six climate factors that influence plant growth for greenhouse climate prediction without any RMSE evaluation.

This research aims to predict the amount of Crude Palm Oil production at PT X using the

LSTM algorithm with an optimal model. The research uses historical data on Crude Palm Oil production from 2019-2023, including the number of Palm Fruit Bunches received and processed from the plantation and outside the plantation. Hopefully, this research will provide deeper insights into the field of Crude Palm Oil production, help companies plan production more efficiently, manage stocks accurately, and support better decision-making in various aspects of oil palm plantation management. The contribution of this research can provide knowledge about predicting palm oil production so that companies can plan more efficient production.

In addition, it is expected that the information obtained from this study can support companies in negotiating better sales contracts, handling equipment maintenance more efficiently, and managing the company's financial cash flow better. Companies are also expected to take appropriate actions to address factors that affect production results. Thus, this study is expected to improve operational efficiency and support more appropriate decision-making in various aspects of oil palm plantation management.

2. FORMULATION OF THE PROBLEM

The problem formulation in this study is how to build a prediction model for palm oil production using Long Short-Term Memory (LSTM).

3. THEORETICAL FRAMEWORK

3.1 Prediction

Prediction is the process of estimating future events based on historical data or existing information, used in various fields to aid decision-making [5]. For example, in business, prediction helps with marketing strategies, production planning, and inventory management. In the financial sector, prediction is used to forecast stock prices, exchange rates, and investment risks. In science and technology, predictions help understand trends in climate change, disease epidemiology, and technological development.

Prediction methods vary from simple statistical techniques to complex machine-learning algorithms. Traditional methods such as linear regression and time series analysis have long been used. However, with the advancement of technology and the availability of big data [6], methods such as Long Short-Term Memory (LSTM) artificial neural networks are gaining popularity. LSTM, a subset of Recurrent Neural Network, effectively handles time series data with complex patterns and long-term relationships.

3.2 Oil Palm

Oil palm (Elaeis guineensis) is a tropical crop that has become an important commodity in the agricultural and edible oil industries. This plant, which originated in West Africa, is now widely grown in tropical regions such as Indonesia, Malaysia, Thailand, and Nigeria [7]. Oil palm fruits produce edible oil used in food, cosmetics, and biodiesel. High productivity rates make oil palm an efficient and popular edible oil crop.

Oil palm grows well in areas with high rainfall and stable temperatures. Oil palm plantations consist of large tracts of land with oil palm trees that start bearing fruit around 3 to 4 years after planting and can be produced for more than 20 years with proper care. The palm oil industry plays an important role in the global economy, especially in major producing countries such as Indonesia and Malaysia, with significant contributions to national income and employment. However, challenges related to environmental sustainability, human rights, and social welfare need to be addressed to ensure sustainable production.

3.3 CPO

CPO or Crude Palm Oil is crude palm oil obtained from the extraction of the flesh of the oil palm fruit, usually from the species Elaeis guineensis, and has not undergone a refining process [8]. Crude palm oil is different from palm kernel oil which is derived from the kernel of the oil palm fruit and coconut oil which is produced from the kernel of the coconut fruit. One of the main differences between crude palm oil is the high content of beta-carotene, which gives it a reddish color and is the initial compound of vitamin A with red or orange pigments in fruits or vegetables.

3.4 Data Mining

Data mining is the process of uncovering valuable patterns or hidden information from large or complex data sets, aiming to discover new knowledge or information to support better decision-making [9]. This process involves several stages: data selection, data pre-processing, modeling, evaluation, and interpretation of results.

One technique in data mining is association analysis, which finds relationships between items in large data sets, such as sales transaction data.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

For example, this technique can show that customers who buy bread also often buy butter. Classification is another technique that predicts new categories based on labeled data, such as predicting whether an email is spam.

Clustering groups data into similar groups based on certain attributes, such as grouping customers based on shopping behavior. Regression models the relationship between dependent and independent variables to make future predictions.

Data mining has wide applications in various fields, including business, science, healthcare, and security. In business, data mining is used for market analysis, risk management, personalized marketing, and process optimization. In science, these techniques help find patterns in experimental or observational data for discoveries or understanding of complex phenomena [10]. Data mining provides valuable insights that support better decisions and innovation in various fields.

3.5 Long Short-Term Memory

LSTM is a type of artificial neural network in the category of Recurrent Neural Networks (RNN) designed to overcome the weakness of RNN in learning long-term dependencies in time series data. Standard RNNs often suffer from the “vanishing gradient” problem which leads to the loss of important information from the beginning of the data sequence. LSTMs overcome this problem by using memory cells and three gates: input, output, and forget. These gates allow the LSTM to retain, update, or delete information selectively, thus capturing long temporal relationships more effectively [11].

The advantages of LSTM make it very useful in various applications involving time series or sequence data, such as weather prediction, natural language processing, sentiment analysis, and stock price prediction. In natural language processing, LSTMs can remember the context of long sentences, aiding machine translation and speech recognition. In stock price prediction, LSTM analyzes market trends and price fluctuation patterns from historical data to provide more accurate predictions [12]. Although it requires significant computational resources, LSTM's ability to handle complex sequence data makes it a superior choice in many machine-learning applications.

The LSTM architecture consists of an input layer, an output layer, and a hidden layer. The LSTM structure consists of memory cells that have three main components: input gate, output gate, and forget gate [13]. The input gate determines the new information stored in the memory cell, the forget gate determines the information that should be deleted, and the output gate controls the information used as output. These gates use a sigmoid neural network to regulate the flow of information in the memory cell.

These gates allow the LSTM to capture and maintain long temporal relationships in time series data. The feedback mechanism in LSTM ensures that important information from the beginning of the data sequence is not lost, making it effective in tasks involving long-term dependencies such as natural language processing and time series prediction.

3.6 Root Mean Squared

Root Mean Squared Error (RMSE) is an evaluation metric that measures the degree of error in a model's prediction results. A lower RMSE value indicates that the model predictions are more accurate. RMSE takes into account the difference between the value predicted by the model and the actual observed value. As such, RMSE helps measure how close the model's estimate is to the true value, and a low RMSE value indicates better prediction performance.

The advantage of RMSE is its ability to provide error values in the same units as the original data, making it easier to interpret. In addition, RMSE is less susceptible to outliers than Mean Squared Error (MSE) because the squared effect on outliers is reduced by rooting. The process of calculating RMSE involves a step of subtracting the actual value from the forecast value, followed by a step of squaring and summing the total squared result. This result is then divided by the amount of data available, and the square root of the total is calculated.

As such, RMSE is an important evaluation tool in evaluating the predictive quality of a model, with lower values signifying better predictive performance.

3.7 Data Normalization

Data normalization methods aim to align the value ranges of several variables to facilitate statistical analysis. One commonly used method is Decimal Scaling normalization, which normalizes the range of values of each attribute into a specific scale by shifting the decimal values of the data in the desired direction. Decimal Scaling normalization has been shown to improve data consistency and balance, especially in the context of classification. In this research, data

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

normalization is important because the LSTM model is sensitive to the scale and distribution of the input data. By using the Decimal Scaling method, each attribute in the dataset has the same range of values, improving the model's performance in predicting palm oil production. Data normalization also helps in accelerating convergence during model training and reduces the risk of overfitting, resulting in more accurate and reliable predictions.

4. MATERIALS AND METHODS PROPOSED

In this section, we provide an overview of the datasets used in this study. In addition, we introduce the prediction approach that we propose, during refinement during the task.

4.1 Dataset

The data collection process in this study involves capturing historical palm oil production data, including the Palm Fruit Bunches received from the farm and off-farm, the total Palm Fruit Bunches processed, and the amount of crude palm oil (CPO) produced daily. Data was obtained from PT X and related institutions, with a total of 1826 data from 2019 to 2023. Data collection was done through interviews, permission from the company, and collection of official reports. Data quality was considered to ensure correctness, accuracy, and completeness. The data was then analyzed and entered into a suitable format for further processing, ensuring the integrity and reliability of the data for palm oil production prediction modeling.

4.2 Proposed Method

In this research, we propose the long short-term memory method where we feel this method is appropriate in the case of time series prediction.

The process of developing a palm oil production prediction model starts with collecting historical production data from 2019 to 2023. The collected data will then go through the Preprocessing stage, where the data will be cleaned from missing or invalid values, as well as normalized if needed. After the data is clean, the next step is to divide the data into two parts, namely training data and testing data. Next, a model will be created using the Long Short-Term Memory (LSTM) algorithm, which is a specialized type of artificial neural network for time series data. The model will be trained using the training data, using Root Mean Square Error (RMSE) as a loss function to measure how well the model predicts palm oil production.

After going through the training stage, the model will be tested using test data that has not previously been seen by the model. The results of the test will be used to calculate the RMSE value, which will give an idea of how close the model prediction is to the actual production value. If the RMSE value has not met the set target, the model will be re-evaluated, and considerations will be made to optimize the model by adjusting the parameters or structure of the LSTM. After a successful evaluation, the model is ready to be used to predict palm oil production for the specified period. The prediction results will provide important insights for planning and decision-making in the palm oil production sector, helping companies to anticipate production fluctuations and optimize operational performance.

5. EXPERIMENT SETUP

In the research setup experiment titled “Application of Data Mining for Palm Oil Production Prediction with Long Short Term Memory Method,” several steps were followed.

5.1 Removing blank values

The stage of removing blank values in the Long Short-Term Memory algorithm is critical to determining the quality of the data that will be used to train and test the model. If there are columns with a high number of blank values, this can be a problem as the missing values can reduce the accuracy and performance of the model. Therefore, the next step is to take appropriate actions, such as deleting rows or columns that contain blank values, replacing blank values with other values, or even finding additional data to fill the blank values. By maintaining consistency in the data and ensuring that empty values are handled properly, we can ensure that the dataset used to train the LSTM model is clean and ready to use. In the author's dataset, there were no empty values found so there was no need to delete the values. Figure 1 of the result of removing blank values is shown below:

```
Total missing values in each column:
Date                 0
Base                 0
TOTAL_TERIMA         0
TOTAL_OLAH           0
JUMLAH_CPO           0
Kebun_Terima         0
Luar Kebun_Terima    0
Kebun_Olah           0
Luar Kebun_Olah      0
dtype: int64
```

Figure 1: Result Of Removing Blank Values

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

5.2 Preprocessing Dataset

5.2.1 Split Dataset

In this step, the palm oil production dataset is divided into two parts: training set and test set. The training set consists of palm oil production data from years before 2023, which is used to train the prediction model. The test set consists of production data starting from 2023, which will be used to test how well the model can predict new data that has never been seen before. With this division, we can ensure that the model we are training has been able to capture the patterns that exist in historical data, as well as evaluate the performance of the model objectively using independent data. In this research, the number of received Palm Fruit Bunches, processed Palm Fruit Bunches and the amount of CPO are predicted, so there are also three splits. The split dataset can be seen in Figures 2, 3, and 4 below:

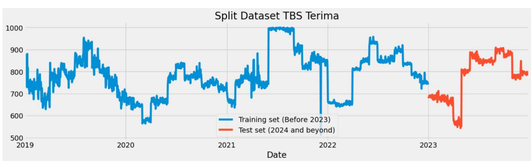

Figure 2: Split Dataset Of Received Palm Fruit Bunches

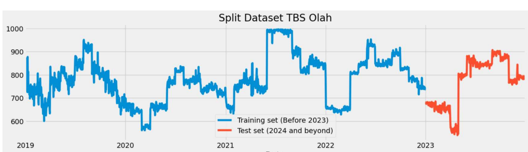

Figure 3: Split Dataset Of Processed Palm Fruit Bunches

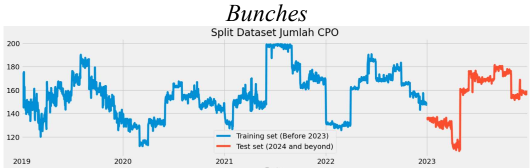

Figure 4: Split Dataset Of Crude Palm Oil Amount

5.3 Data Conversion

In this step, the 'date' column in the dataset was converted to the DateTime data type using the pd.to_datetime() function. The datetime data type is beneficial as it allows us to easily work with dates and times in data analysis. After that, the 'Date' column is set as the index of the DataFrame using the set_index('Date', inplace=True) function. By setting the 'Date' column as an index, we can easily access the data by date, which will be very useful when performing time-based data analysis, such as forecasting or trend analysis. The dataset conversion can be seen in Figure 5 below:

```
# Untuk mengonversi kolom 'tanggal' menjadi tipe data datetime
dataset['Date'] = pd.to_datetime(dataset['Date'])

# mengatur kolom 'Date' sebagai indeks DataFrame
dataset.set_index('Date', inplace=True)
```

Figure 5: Dataset conversion

5.4 Normalization

Normalization in LSTM is to transform the range of input values into a smaller range, generally between 0 to 1, to better match the activation characteristics of the sigmoid function commonly used in LSTM. The sigmoid activation operates more efficiently when the inputs are in a range close to 0 to 1. In this step, the dataset is normalized using Min-Max Scaling. Examples of normalized datasets can be seen in Tables 1, 2, and 3 below:

Table 1: Example Of Normalized Received Palm Fruitbunches Dataset

Receive Palm Fruit Bunches dataset before normalization	***Receive Palm Fruit Bunches dataset after normalization***
773	0.562
853	0.875
820	0.746
837	0.812
790	0.628
876	0.964
778	0.582
729	0.390
880	0.980
754	0.488

Table 2: Example Of Normalized Olive Palm Fruit Bunches Dataset

Olive Palm Fruit Bunches dataset before normalization	***Olive Palm Fruit Bunches dataset after normalization***
770	0.604
850	0.889
816	0.768
833	0.829
785	0.658
871	0.964
763	0.580
725	0.444
877	0.985
750	0.533

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

Table 3: Example Of Normalized CPO Quantity Dataset

Dataset Number of CPO before normalization	Dataset Number of CPO after normalization
154	0.605
170	0.889
163.2	0.764
166.6	0.832
157	0.658
174.2	0.942
152.6	0.580
145	0.444
175.4	0.948
150	0.513

5.5 LSTM Model Training

This stage is the Long Short-Term Memory (LSTM) model training preparation stage, which focuses on the formation of a suitable training dataset for the LSTM model. This process aims to create input data (X_train) and output targets (y_train) that are used by the LSTM model to learn and understand the time sequence patterns in the normalized data. This stage is important to ensure that the LSTM model can learn the temporal relationships and patterns present in the time sequence data. By forming data pairs (input-output) of a time window consisting of the previous 60 days and the value to be predicted on day 61, the model can understand and predict future values based on patterns found in historical data. The training of this model can be seen in Figures 6, 7, and 8 below:

```
# Persiapan pelatihan model
X_train = []
y_train = []
for i in range(60, len(training_set_scaled)):
    X_train.append(training_set_scaled[i-60:i, 0])
    y_train.append(training_set_scaled[i, 0])
X_train, y_train = np.array(X_train), np.array(y_train)
```

Figure 6: Training Code Of The Receiving Palm Fruit Bunches Model

```
X_train2 = []
y_train2= []
for i in range(60, len(training_set_scaled2)):
    X_train2.append(training_set_scaled2[i-60:i, 0])
    y_train2.append(training_set_scaled2[i, 0])
X_train2, y_train2 = np.array(X_train2), np.array(y_train2)
```

Figure 7: Model Training Score Of Palm Fruit Bunches Processed

```
X_train3 = []
y_train3= []
for i in range(60, len(training_set_scaled3)):
    X_train3.append(training_set_scaled3[i-60:i, 0])
    y_train3.append(training_set_scaled3[i, 0])
X_train3, y_train3 = np.array(X_train3), np.array(y_train3)
```

Figure 8: Cpo Quantity Model Training Score

5.6 Test Data Preparation

Test data preparation is done by combining the historical data used for training with the test data. Historical data was selected as the length of the test dataset plus the previous 60 entries, as the LSTM model requires several historical values to make predictions. The chosen data was then reshaped into a two-dimensional array with one column, and normalized using the same scaler as the training data. This step ensures that the LSTM model can use relevant historical data to make accurate predictions. By preparing the test data consistently with the training process, the accuracy of the model's prediction of palm oil production can be improved. The training test data can be seen in Figures 9, 10, and 11 below:

```
dataset_total = pd.concat((dataset["TOTAL_TERIMA"][:'2022'],dataset["TOTAL_TERIMA"]['2023':]),axis=0)
inputs = dataset_total[len(dataset_total)-len(test_set) - 60:].values
inputs = inputs.reshape(-1,1)
inputs  = sc.transform(inputs)
```

Figure 9: Training Data For The Receiving Palm Fruit Bunches Test

```
dataset_total2 = pd.concat((dataset["TOTAL_OLAH"][:'2022'],dataset["TOTAL_OLAH"]['2023':]),axis=0)
inputs2 = dataset_total2[len(dataset_total2)-len(test_set2) - 60:].values
inputs2 = inputs2.reshape(-1,1)
inputs2  = sc2.transform(inputs2)
```

Figure 10: Training Test Data Of Processed Palm Fruit Bunches

```
dataset_total3 = pd.concat((dataset["JUMLAH_CPO"][:'2022'],dataset["JUMLAH_CPO"]['2023':]),axis=0)
inputs3 = dataset_total3[len(dataset_total3)-len(test_set2) - 60:].values
inputs3 = inputs3.reshape(-1,1)
inputs3  = sc3.transform(inputs3)
```

Figure 11: Training Of CPO Quantity Test Data

We created a list X_test containing input data for the model from the test dataset, where each entry is a 60-day sequence of historical data. The data in X_test was then reshaped into a three-dimensional array to meet the input format expected by the LSTM model. After that, the LSTM model is trained to predict palm oil production values based on the given input data. The predictions generated by the model are still in a normalized form, so the author uses sc.inverse_transform to return the prediction scale to its original scale. Thus, the predicted value is on the same scale as the original palm oil production data. The test data preparation code can be seen in Figures 12, 13, and 14 below:

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

```
X_test = []
for i in range(60, len(inputs)):
    X_test.append(inputs[i-60:i, 0])
X_test = np.array(X_test)
X_test = np.reshape(X_test, (X_test.shape[0], X_test.shape[1], 1))
predicted_stock_price = regressor.predict(X_test)
predicted_stock_price = sc.inverse_transform(predicted_stock_price
```

Figure 12: Test Data Preparation Of Accepted Palm Fruit Bunches

```
X_test2 = []
for i in range(60, len(inputs2)):
    X_test2.append(inputs2[i-60:i, 0])
X_test2 = np.array(X_test2)
X_test2 = np.reshape(X_test2, (X_test2.shape[0], X_test2.shape[1], 1))
predicted_stock_price2 = regressor2.predict(X_test2)
predicted_stock_price2 = sc2.inverse_transform(predicted_stock_price2)
```

Figure 13: Preparation Of Test Data On Processed Palm Fruit Bunches

```
X_test3 = []
for i in range(60, len(inputs3)):
    X_test3.append(inputs3[i-60:i, 0])
X_test3 = np.array(X_test3)
X_test3 = np.reshape(X_test3, (X_test3.shape[0], X_test3.shape[1], 1
predicted_stock_price3 = regressor3.predict(X_test3)
predicted_stock_price3 = sc3.inverse_transform(predicted_stock_price
```

Figure 14: Preparation Of Test Data For Total CPO

5.7 LSTM Implementation

The LSTM architecture used in this study consists of multiple layers designed to predict palm oil production with high accuracy. The model starts by creating a Sequential object from Keras, which is used to build the layered neural network. In this study, the authors also performed fine-tuning using the following hyperparameters, which were taken based on recommendations for LSTM:

1. Epoch: 50
2. Batch size: 32

This study uses a four-layer LSTM architecture to predict palm oil production. The use of multiple layers of LSTM allows the model to capture complex temporal information from historical palm oil production data. To prevent overfitting, the authors apply a 20% dropout rate to each LSTM layer. The last layer is a dense layer with 1 unit to produce palm oil production prediction output.

After building the model architecture, the model was compiled using the optimizer 'RMSprop' and the loss function 'mean_squared_error'. The training process was conducted 4 times with various configurations, including 50 and 100 epochs with batch sizes of 32 and 64. The results of the training process can be seen in Table 4 below:

Table 4: Training Experiments

50	32	0.0225	0.1383
100	32	0.0257	0.1381
50	64	0.0383	0.1279
100	64	0.0298	0.1707

In the table, it can be seen that from 4 experiments, the best results were obtained at 50 epochs and 32 batch sizes. Based on these experiments, the author chose the number of epochs and batch size in this study. This setting helps the model to learn from the data gradually and refine its weights to improve prediction accuracy.

5.8 Evaluation

In this study, an evaluation of the model that has been made is needed, the evaluation stage using Root Mean Squared Error (RMSE) aims to measure how well the LSTM model has been trained in predicting palm oil production. This assessment is important because it gives an idea of how close the predicted value produced by the model is to the actual value of the data that has never been seen before. The RMSE evaluation results for palm oil production prediction are shown in Figures 15, 16, and 17 below:

```
return_rmse(test_set_scaled1,predicted_stock_price_d)

RMSE : 0.12384899092714552.
```

Figure 15: Evaluation Results Of Accepted Palm Fruit Bunches

```
return_rmse(test_set_scaled2,predicted_stock_price_d)

RMSE : 0.11770094237784422.
```

Figure 16: Evaluation Result Of Palm Fruit Bunches Processed

```
return_rmse(test_set_scaled3,predicted_stock_price_d)

RMSE : 0.11770094237784419.
```

Figure 17: Total Cpo Evaluation Results

5.9 PREDICTION RESULTS

The prediction results for the next three-year period can be seen in Figures 18, 19, and 20 below:

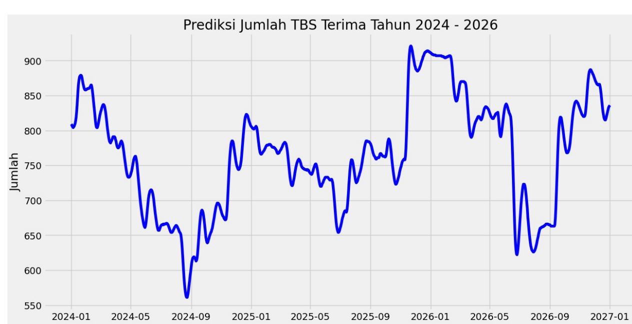

Figure 18: Visualization Of Prediction Results Of Palm Fruit Bunches Received in 2024 – 2026

Journal of Theoretical and Applied Information Technology
15th November 2024. Vol.102. No. 21

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

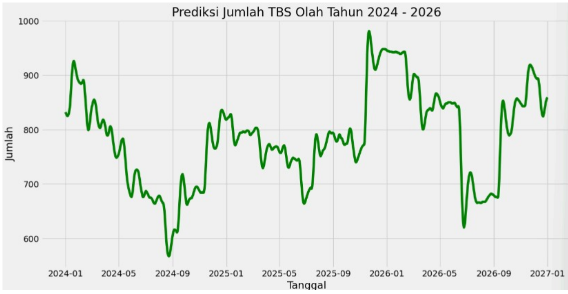

Figure 19: Visualization Of Prediction Results Of Palm Fruit Bunches Processed In 2024 – 2026

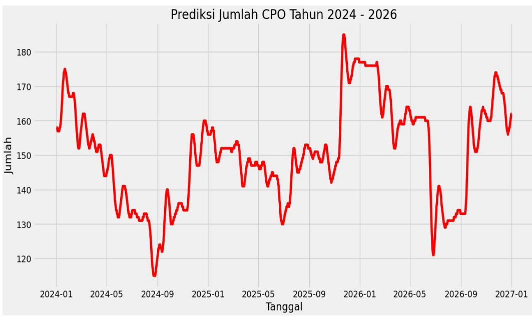

Figure 20: Visualization Of Prediction Results Of CPO In 2024 – 2026

5.10. Difference from Prior Work

Research by Anwar surveyed short-term highway traffic forecasting algorithms based on short-term memory (LSTM) deep learning models. However, this study only focused on six climate factors affecting plant growth for greenhouse climate prediction without any RMSE evaluation. In this study, we conducted RMSE evaluation with results of 0.1360 for total FFB received, 0.1279 for total FFB processed, and 0.1279 for total CPO indicating good potential.

Research by Yuwen proposed a greenhouse climate prediction model. The focus of this research is on six climate factors affecting plant growth, including temperature, humidity, lighting, carbon dioxide concentration, soil temperature, and soil moisture, and promotes the GCP_lstm model for greenhouse climate prediction. However, this study only focuses on six climate factors that influence plant growth for greenhouse climate prediction without any RMSE evaluation in this study, we conducted RMSE evaluation with results of 0.1360 for total FFB received, 0.1279 for total FFB processed, and 0.1279 for total CPO indicating good potential.

6. CONCLUSIONS

After designing, testing the system, and analyzing it, it was concluded that, among others, the system built was able to predict the amount of palm oil production well. Testing the Long Short-Term Memory method using Hyperparameters for 50 epochs and a batch size of 32 produced good prediction values for palm oil production. The evaluation results using the RMSE value with the RMSE value for the total FFB received was 0.1238, the total FFB processed was 0.1177, and the total CPO was 0.1177. Where these results indicate that the LSTM model has a good level of accuracy in predicting palm oil production. The contribution of this research can provide knowledge about predicting palm oil production so that companies can plan more efficient production.

AUTHOR CONTRIBUTION

Handrizal: Conceptualization, Validation, data analysis refined methodology and supervised overall process of research.
Dewi Sartika Ginting: Conceptualization, Methodology, helped review and edit the manuscript.
Herriyance: Conceptualization, Formal analysis and investigation, Algorithms, and data analysis
Syabrina Ramadhani Kamal: Programmer, Development and overall implementation, and preliminary writing.

REFERENCES

[1] Kamath, P., Patil, P., S, S., Sushma, & S, S. (2021). Crop yield forecasting using data mining. Global Transitions Proceedings, 2(2), 402–407. https://doi.org/10.1016/j.gltp.2021.08.008
[2] Song, X., Liu, Y., Xue, L., Wang, J., Zhang, J., Wang, J., Jiang, L., & Cheng, Z. (2020). Time-series well performance prediction based on Long Short-Term Memory (LSTM) neural network model. Journal of Petroleum Science and Engineering, 186, 106682. https://doi.org/10.1016/j.petrol.2019.10668
[3] A. Khan, M. M. Fouda, D. -T. Do, A. Almaleh and A. U. Rahman, "Short-Term Traffic Prediction Using Deep Learning Long Short-Term Memory: Taxonomy, Applications, Challenges, and Future Trends," in IEEE Access, vol. 11, pp. 94371-94391, 2023, doi: 10.1109/ACCESS.2023.3309601
[4] Liu Y, Li D, Wan S, et al. A long short-term memory-based model for greenhouse climate prediction. Int J Intell Syst. 2022; 37: 135-151. https://doi.org/10.1002/int.22620
[5] Wang, L., Li, J., Guo, S., Xie, N., Yao, L., Cao, Y., ... & Sun, D. (2020). Real-time estimation and prediction of mortality caused by COVID-19 with patient information-based algorithm. Science of the total environment, 727, 138394.

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

[6] Barrera-Animas, A. Y., Oyedele, L. O., Bilal, M., Akinosho, T. D., Delgado, J. M. D., & Akanbi, L. A. (2022). Rainfall prediction: A comparative analysis of modern machine learning algorithms for time-series forecasting. Machine Learning with Applications, 7, 100204.
[7] Ilyas, R. A., Sapuan, S. M., Ibrahim, M. S., Wondi, M. H., Norrrahim, M. N. F., Harussani, M. M., ... & Haris, N. I. N. (2022). Introduction to oil palm biomass. In Oil Palm Biomass for Composite Panels (pp. 3-38). Elsevier.
[8] Nizam, A. F. A., & Mahmud, M. S. (2021). Food quality assurance of crude palm oil: a review on toxic ester feedstock. OCL, 28, 23.
[9] Gupta, M. K., & Chandra, P. (2020). A comprehensive survey of data mining. International Journal of Information Technology, 12(4), 1243-1257.
[10] Sunhare, P., Chowdhary, R. R., & Chattopadhyay, M. K. (2022). Internet of things and data mining: An application-oriented survey. Journal of King Saud University-Computer and Information Sciences, 34(6), 3569-3590.
[11] Okut, H. (2021). Deep learning for subtyping and prediction of diseases: Long-short term memory. Deep Learning Applications.
[12] Nourbakhsh, Z., & Habibi, N. (2023). Combining LSTM and CNN methods and fundamental analysis for stock price trend prediction. Multimedia Tools and Applications, 82(12), 17769-17799.
[13] Xia, K., Huang, J., & Wang, H. (2020). LSTM-CNN architecture for human activity recognition. IEEE Access, 8, 56855-56866.