

Predictive Analysis of Raw Material Stock at Puri Food and Healthy, an SME, Using the Long Short-Term Memory (LSTM) Method

Palma Juanta^{[1]*}, Pery Chandria Bangun^[2], Muhammad Fachrul Nizam^[3]

Department of Information System, Faculty of Science and Technology^{[1], [2], [3]}

Universitas Prima Indonesia

Medan, Indonesia

palmajuanta@unprimdn.ac.id^[1], ferybangun444@gmail.com^[2], fahrolnizam63@gmail.com^[3]

Abstract— Micro, Small, and Medium Enterprises play a vital role in the Indonesian economy, yet face significant challenges in managing raw material inventories, particularly for perishable commodities such as coconut sap (nira). This study applies and optimizes the Long Short-Term Memory (LSTM) method to predict raw material stock levels for coconut sap at Puri Food and Healthy, an SME, using five years of daily historical data (January 1, 2020–December 31, 2024; 2,191 entries). A descriptive and experimental quantitative approach was employed to develop a deep learning-based predictive model, with data obtained through inventory documentation and interviews with SME managers. The research process encompassed data preparation, collection, normalization, LSTM model construction using Python and TensorFlow in Google Colab, and evaluation using Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R² Score. Results show the model achieved an MAE of 5.31 and an RMSE of 6.94, indicating moderate prediction error. However, the R² value of 0.0711 suggests very low explanatory power, potentially due to underfitting or data limitations. Notably, multi-step forecasting was applied to generate projections for 2026–2027 despite having historical data only through 2024, with these extended forecasts intended as experimental. The model successfully learned seasonal patterns but requires further optimization to improve predictive accuracy. This study advances AI-based inventory management for SMEs, supporting operational efficiency, waste reduction, and risk mitigation in raw material supply chains.

Keywords— Stock Prediction, SME, LSTM, Nira, Deep Learning

I. INTRODUCTION

Indonesia's economy is built upon micro, small, and medium-sized businesses, which employ over 97% of the labor force and account for more than 60% of the country's GDP [1][2]. Despite their vital role, SMEs continue to struggle with supply chain and inventory management issues, including demand fluctuations, supply delays, insufficient storage space, and financial constraints [3].

Inventory management is even more crucial for commodities like coconut sap (nira), which are prone to rapid deterioration in quality and economic value. Because it is seasonal, extremely perishable, and susceptible to storage

conditions, Nira needs predictive methods that can respond to time-dependent trends [4]. A stock management system that can anticipate changes in the demand for raw materials is necessary for Nira processing SME, such as Puri Food and Healthy.

Because it still relies on manual inventory recording, Puri Food and Healthy, an SME that processes nira into palm sugar products, has a slow decision-making process [5]. Digitizing inventory systems may improve operational efficiency and logistics [6]. Predictive approaches, such as the LSTM model, are well-suited for these situations because they can learn temporal patterns from historical data [7].

This approach has been used extensively across a variety of forecasting scenarios, including predicting the number of SMEs [8], the prices of agricultural commodities [9], the price of crude oil [10], and the supply of food products like Sari Roti [11]. Kusuma et al.'s study demonstrated that LSTM can forecast bread product inventories with an MAPE accuracy of less than 10% [11]. Additionally, Susilo et al. [12] found that LSTM performed well in logistics distribution, achieving a MAPE of about 3.6%. This reinforces LSTM's status as an adaptable model for handling complex time-series data, such as raw material stock data.

The adoption of AI in micro-scale SMEs remains low [16], and traditional techniques such as decision tree algorithms struggle to detect complex patterns [13][15]. As a result, the goal of this study is to develop and refine an LSTM-based model for predicting NIRA stock levels at Puri Food and Healthy, assess its performance using MAE, RMSE, and R², and investigate its potential to enhance inventory management practices.

The purpose of this research is to apply and improve the model for predicting Nira raw material inventories at Puri Food and Healthy SME by evaluating it using a range of performance indicators, including Root Mean Square Error (RMSE), Mean Absolute Error (MAE), and R² Score. Given its extreme perishability and seasonality, Nira needs a predictive strategy that can account for time-based dynamic trends [17]. As a result, this study not only addresses the practical needs of SMEs but also advances the use of deep learning in the management of SME supply chains.

II. RESEARCH METHODOLOGY

This study falls under the quantitative research category and employs both descriptive and experimental approaches. In Puri Food and Healthy SME, the current state of affairs about the management of nira raw material inventories is described using the descriptive method. Using the Long-Short Term Memory technique, the experimental approach is used to create and evaluate a prediction model for nira raw material inventories. Since the purpose of this sort of study is to identify viable remedies for the issue of controlling nira raw material inventories at Puri Food and Healthy SME, it is applicative.

The study's population consists of Puri Food and Healthy SMEs, businesses in the food and beverage sector that use natural raw materials, particularly nira as their primary ingredient.

The study's sample was selected using the purposive sampling method, with specific criteria. The sample selection criteria include SME that uses nira as its primary raw material, produces food and drinks that are made with natural ingredients, keeps thorough historical records of its raw material inventories, and is prepared to take part in interviews and surveys. Because they satisfy all these requirements and have a stock recording system that can be analyzed using the LSTM technique, Puri Food and Healthy SME are used as the primary sample in this study.

The study started by identifying and gathering historical information on Nira stocks, creating questionnaires, and making technical preparations. Quantitative data was collected through documentation, whereas qualitative data was collected through interviews and surveys with SME managers. The information was then processed and analyzed using the LSTM prediction model. A final report, which included analysis, results, and research recommendations, was then used to evaluate the predictive outcomes and determine the model's efficacy.

A. Data Preparation

The first step in the research preparation phase is to identify issues and gather historical information pertaining to nira raw material stocks at Puri Food and Healthy SME in order to serve as a basis for analysis. The purpose of this stage is to determine the pattern of stock availability and any challenges that may arise in its management. Additionally, interview schedules were created to delve further into the stock recording and management process, and research tools, such as questionnaires, were developed to gather corroborating data from SME managers.

B. Data Collection

Quantitative data collection was conducted during the data collection phase, with historical records of Nira raw material inventories from Puri Food and Healthy SME serving as the primary basis for predictive analysis. Additionally, comprehensive interviews with SME managers and employees were conducted to collect qualitative data on operational practices, inventory management techniques, and issues encountered.

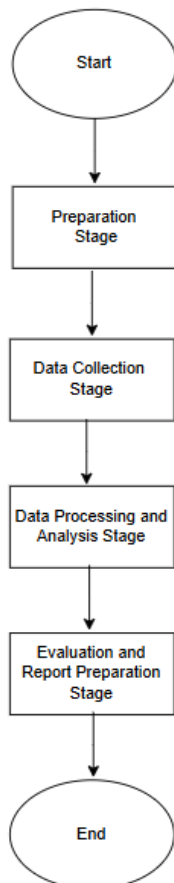


Figure 1. Flowchart of Nira Raw Material Stock Prediction

TABLE 1. DATASET

Date	Stock
2020-01-01	52
2020-01-02	27
2020-01-03	24
2020-01-04	25
2020-01-05	30
2020-01-06	32
2020-01-07	39
2020-01-08	24
2020-01-09	36
2020-01-10	39
2020-01-11	24
2020-01-12	40
2020-01-13	29
2020-01-14	21
2020-01-15	21
2020-01-16	34
2020-01-17	30
2020-01-18	25
2020-01-19	41
2020-01-20	28
2020-01-21	40
.....
2024-12-31	38

A sample of historical daily stock data from January 1, 2020, to December 31, 2024, is shown in Table 1. The dataset consists of two columns: 'date', which shows the time the recording was made, and 'stock', which shows the amount of stock for that day. This information serves as the basis for analysis and predictive modeling in the LSTM method. The Dataset Description and Statistical Distribution utilized in this research are explained below.

1. Data Type and Time Period

- The data used is daily time series data on raw material stock at Puri Food and Healthy.
- Observation period: January 1, 2020 – December 31, 2024 (total of 2,191 entries).
- Available variables:
 - date (date of recording)
 - stock (stock quantity on that day, in units/pieces).

2. Stock Value Distribution per Year

- 2020: Initial production fluctuations with seasonal trends following the harvest period.
- 2021: Tended to be stable around the previous year's average.
- 2022: The lowest stock quantity recorded, totaling 11,794 units.

- 2023: Highest stock level recorded, totaling 13,587 units.
- 2024: Experienced a slight decrease compared to the previous year, totaling 13,364 units.

3. Seasonal Pattern

- A seasonal pattern is evident, where stock increases during specific periods (likely the harvest season) and decreases during periods outside the harvest season.
- These fluctuations are consistent every year, indicating the presence of periodic factors that influence stock levels.

4. Daily Distribution and Data Dispersion

- Daily stock values fall within a specific range (e.g., between 20–80 units per day when divided into incoming/outgoing stock).
- The distribution shows significant peak seasons and off-seasons.

5. Data Quality

- There are no missing values, so the data is suitable for modeling after normalization.
- The date format is consistent (datetime), and stock values are integers.

C. Data Processing and Analysis

To ensure data quality and readiness for computational analysis during the data processing and analysis phase, all acquired data undergo cleaning, normalization, and coding. After the historical data on prepared nira raw material stocks has been properly organized, a predictive model based on LSTM is created and trained. To produce accurate stock predictions and support sound inventory management in SMEs, this model analyzes seasonal trends and data patterns.

D. Data Evaluation

To assess the LSTM model's accuracy and capacity to predict NIRA raw material stocks, the evaluation phase includes a comprehensive examination of its predictions. Performance measures such as RMSE, MAE, and R2 Score are used to evaluate how well anticipated values and actual data match up. In addition, a comprehensive data analysis is performed at this stage, the main conclusions of the study findings are presented, and strategic recommendations are made that SMEs can use in data-based stock management.

III. RESULTS AND DISCUSSION

At this point, the findings of the data analysis are presented, along with a discussion of the research's conclusions. The outcomes presented here include the model training method, an evaluation of its predictive accuracy, and an interpretation of the analyzed data. The discussion aims to determine the accuracy of the LSTM model in predicting the quantity of sap raw materials and the implications of these predictions for inventory management at Puri Food and Healthy SME.

A. Data Normalization

Before inputting the data into the LSTM model, several preprocessing steps are performed. The initial step is data

cleaning, which involves eliminating missing values and correcting data replication.

```
[21] # Normalisasi data
scaler = MinMaxScaler()
data_scaled = scaler.fit_transform(data_stok)
```

Figure 2. Data Normalization

The figure above shows that the data is normalized using the MinMaxScaler method to ensure all values fall between 0 and 1, which helps speed up the training process and avoid the dominance of high values. Following this, the data is transformed into a windowed dataset, also known as a sliding window.

B. Data Sharing and Training

For ease of data sharing, the historical raw material stock data is split into training data (training data) and testing data (testing data). The purpose of this division is to train the prediction model on a subset of the data and then test its accuracy on previously unseen data.

```
[29] # Split train dan test
from sklearn.model_selection import train_test_split
X_train, X_test, y_train, y_test = train_test_split(X, y, test_size=0.2, random_state=42)
```

Figure 3. Data Sharing and Training

The data is divided in this investigation with a ratio of 80% for training and 20% for testing, as shown in the diagram above. Since the data are a time series, the division procedure is performed without randomization (shuffle=False) to preserve the data's chronological order. It's crucial to prevent data leakage and ensure the model learns from historical trends in a logical, time-ordered manner, as they occur in the real world.

C. Model Architecture

The LSTM model was built in Google Colab using Python, TensorFlow, and Keras.

```
[37] # Import library
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt
from sklearn.preprocessing import MinMaxScaler
from sklearn.metrics import mean_absolute_error, root_mean_squared_error, r2_score
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense
import math
from sklearn.model_selection import train_test_split
from tensorflow.keras.models import Sequential
from tensorflow.keras.layers import LSTM, Dense
from tensorflow.keras.callbacks import History
```

Figure 4. Import Library

The Adam optimizer, which is frequently used for time series data due to its rapid convergence, and the Mean Squared Error (MSE) loss function were used to train the model and prevent overfitting. Additionally, a dropout layer with a dropout value of 0.2 was included.

```
[40] # Model Architecture
model = Sequential()
model.add(LSTM(64, return_sequences=True, input_shape=(10, 1)))
model.add(LSTM(32, return_sequences=False))
model.add(Dense(1))
model.compile(optimizer='adam', loss='mse')
```

Figure 5. Model Architecture

This research employs a two-layer, sequential LSTM model architecture. With the parameter return_sequences=True, the first layer uses 64 memory units, allowing the output of the entire time step to be sent to the next layer. The second layer has 32 LSTM units without return sequences. The model is then finished with a single-neuron dense output layer that produces a single prediction. The model was constructed using the Mean Squared Error (MSE) loss function and Adam's optimizer.

D. Model Training

```
history = model.fit(X_train, y_train, validation_data=(X_test, y_test),
epochs=30, batch_size=16, verbose=1)
```

```
Epoch 1/30
107/107 ----- 8s 49ms/step - loss: 0.0054 - val_loss: 0.0048
Epoch 2/30
107/107 ----- 10s 44ms/step - loss: 0.0036 - val_loss: 0.0048
Epoch 3/30
107/107 ----- 5s 48ms/step - loss: 0.0045 - val_loss: 0.0049
Epoch 4/30
107/107 ----- 10s 42ms/step - loss: 0.0038 - val_loss: 0.0047
Epoch 5/30
107/107 ----- 6s 51ms/step - loss: 0.0035 - val_loss: 0.0050
Epoch 6/30
107/107 ----- 4s 41ms/step - loss: 0.0042 - val_loss: 0.0049
Epoch 7/30
107/107 ----- 5s 48ms/step - loss: 0.0035 - val_loss: 0.0048
Epoch 8/30
107/107 ----- 5s 43ms/step - loss: 0.0035 - val_loss: 0.0047
Epoch 9/30
107/107 ----- 5s 41ms/step - loss: 0.0035 - val_loss: 0.0050
Epoch 10/30
107/107 ----- 5s 50ms/step - loss: 0.0034 - val_loss: 0.0047
Epoch 11/30
107/107 ----- 10s 45ms/step - loss: 0.0043 - val_loss: 0.0048
Epoch 12/30
107/107 ----- 5s 45ms/step - loss: 0.0034 - val_loss: 0.0048
Epoch 13/30
107/107 ----- 5s 42ms/step - loss: 0.0038 - val_loss: 0.0050
Epoch 14/30
107/107 ----- 6s 50ms/step - loss: 0.0036 - val_loss: 0.0047
Epoch 15/30
107/107 ----- 10s 50ms/step - loss: 0.0040 - val_loss: 0.0048
Epoch 16/30
107/107 ----- 9s 42ms/step - loss: 0.0037 - val_loss: 0.0051
Epoch 17/30
107/107 ----- 6s 51ms/step - loss: 0.0050 - val_loss: 0.0049
Epoch 18/30
107/107 ----- 10s 50ms/step - loss: 0.0036 - val_loss: 0.0047
```

Figure 6. Model Training

The model was trained for 30 epochs with a batch size of 16 using the model.fit() function using training data (X_train, y_train) and validation data (X_test, y_test). The loss (error on the training data) and val_loss (error on the validation data) were recorded at each epoch during training. The val_loss ranged from 0.0051 to 0.0054, while the loss values ranged from 0.0033 to 0.0043, according to the training results. The somewhat consistent loss and val_loss values demonstrate the model's ability to learn from the training data without becoming overly overfitted. Consequently, the model can anticipate outcomes and accurately understand data patterns.

E. Loss Training Model

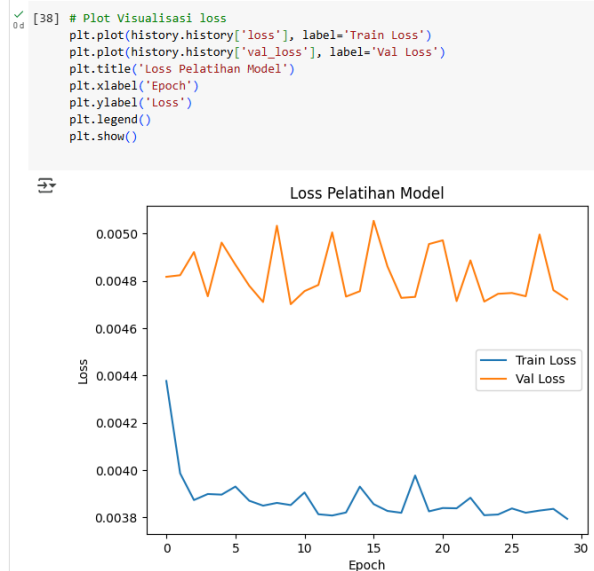


Figure 7. Loss Training Model

The progression of the error values on the training and validation data throughout the training process can be observed by visualizing the loss graph. Based on the comparison between loss and val_loss, this graph shows that the model undergoes a steady learning process and helps identify possible over- or underfitting.

F. Load and Display Data

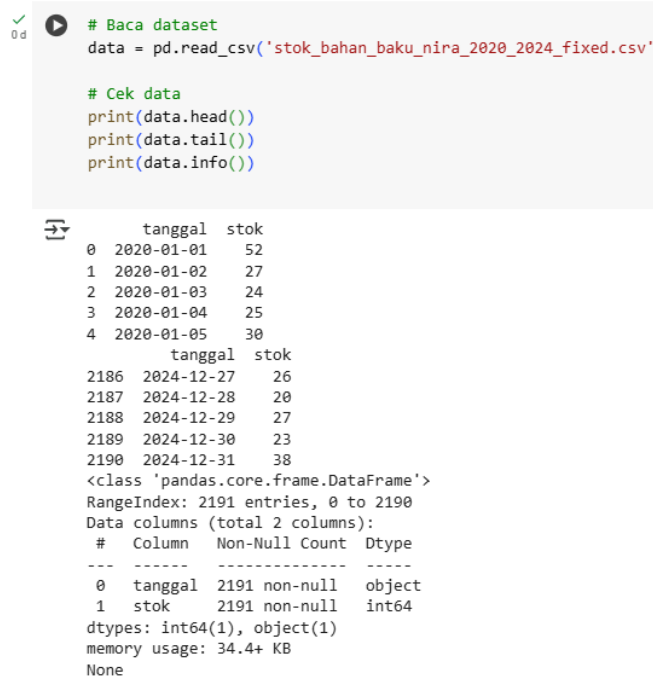


Figure 8. Displaying Data Per Day

This analysis is based on a dataset of 2,191 records that capture daily information on the stock of Nira raw materials at Puri Food and Healthy SME from January 1, 2020, to December 31, 2024. There are two variables in this dataset: 'date,' which is of object data type (datetime), and 'stock,' which

is of numeric type (integer). Because there are no missing values, all data is thoroughly captured, making it appropriate for usage in the pre-processing and predictive modeling phases.

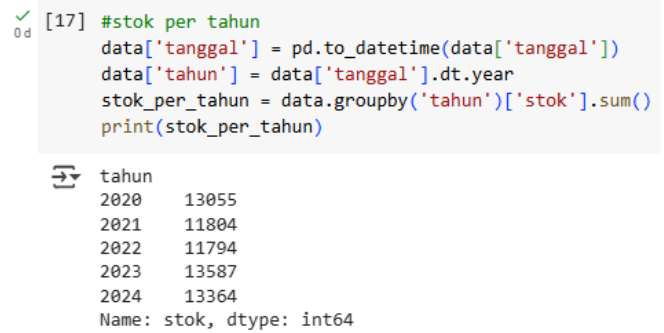


Figure 9. Displaying Data Per Year

Based on the results of the recapitulation of the total stock of nira raw materials at Puri Food and Healthy SME during the period 2020 to 2024, there are fluctuations in the amount of stock each year. The year 2023 recorded the highest total stock of 13,587 units, followed by 2024 with 13,364 units. Meanwhile, the lowest stock amount occurred in 2022 at 11,794 units. This annual variation may be related to dynamics in the production and distribution of raw materials, as suggested in previous studies [4][17], which indicate that factors such as weather conditions, harvest patterns, and market demand often influence the availability of perishable commodities like nira. To make sound strategic decisions, particularly in future stock planning and supply chain management, it is essential to build on these findings.

G. Model Evaluation

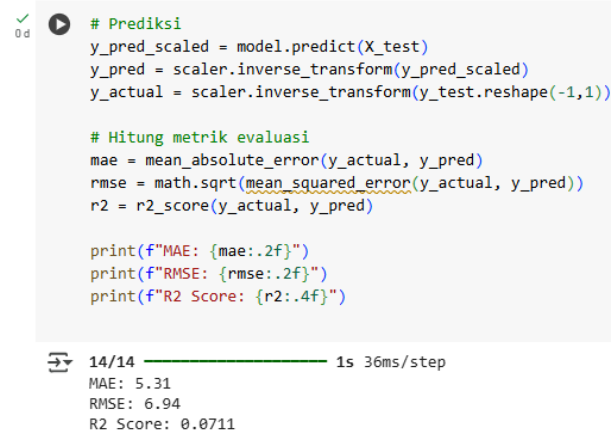


Figure 10. Model Evaluation

The model is evaluated using three indicators: the Mean Absolute Error (MAE), the Root Mean Squared Error (RMSE), and the R-squared (R2). The model's average prediction error against actual data is shown with an RMSE of 6.94 and an MAE of 5.31. On the other hand, the model's R2 value of 0.0711 indicates that it still has a limited ability to account for the variability in the target data.

H. Visualization of Prediction Results

Visualization is used to compare the model's predictions to the real data. The resulting graph shows that the LSTM model's predictions closely match the actual data trends. Seasonal trends, such as an increase in inventory in certain months and a decline during times of reduced demand, were accurately predicted by the model. This further supports the model's capacity to both reflect historical trends and use them to forecast future values.

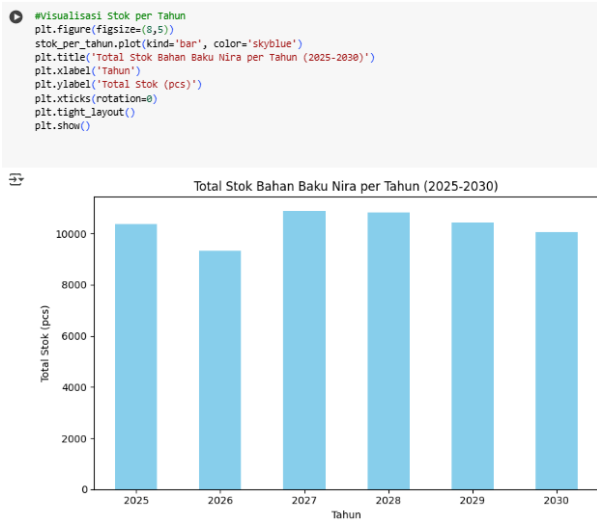


Figure 11. Visualization of Prediction Results

Based on the available data, the total product stock fluctuated during the period 2025 to 2030. In 2025, the total stock was recorded at around 10,506 pcs, then decreased to around 9,211 pcs in 2026. However, in 2027, there was a significant increase, reaching around 11,252 pcs. This number declined slightly in 2028 to around 11,145 pcs and again in 2029 to around 10,512 pcs. In 2030, the total stock was approximately 10,103 pcs. The highest stock was recorded in 2027 at 11,252 pcs, while the lowest stock occurred in 2026 at 9,211 pcs.

Overall, this data shows year-to-year variations in stock levels, influenced by factors such as market demand, production strategies, and distribution efficiency.

IV. CONCLUSION

The study's results imply that the Long Short-Term Memory (LSTM) model was successfully used to predict the inventory of nira raw materials at Puri Food and Healthy SME. With an acceptable degree of error, this deep learning model successfully learned temporal patterns from five years of historical stock data (2020–2024) and made predictions. The model's steady performance during training was evident in the consistently low, stable values of loss and val_loss. The evaluation, which used performance metrics, showed a respectable level of predictive accuracy, with a Root Mean

Square Error (RMSE) of 6.94 and a Mean Absolute Error (MAE) of 5.31. The model's relatively low R2 value of 0.0711, however, indicates that it still cannot fully account for and explain the data's variation. This might be attributed to limited predictive features, the omission of external variables (such as weather conditions, harvest cycles, and market demand), potential underfitting of the model architecture, and the use of multi-step forecasting, which can exacerbate cumulative prediction errors. Additionally, implementing this forecasting model can support more efficient procurement planning, optimize inventory management, and help minimize financial losses from raw material waste, with the highest predicted stock in 2027 at 11,252 pcs and the lowest in 2026 at 9,211 pcs.

REFERENCES

- [1] Kementerian Koperasi dan UKM Republik Indonesia, "Data UMKM Tahun 2023," Jakarta, 2023.
- [2] Kementerian Koperasi dan UKM Republik Indonesia, Laporan Tahunan UMKM Indonesia, Jakarta, 2023.
- [3] A. Wibowo, "Manajemen Rantai Pasok UMKM di Era Digital," Jurnal Logistik dan Rantai Pasok, vol. 7, no. 2, pp. 115–123, 2021.
- [4] R. Sari dan B. Nugroho, "Analisis Daya Simpan dan Mutu Nira dalam Pengolahan Gula Merah," Jurnal Teknologi Pangan, vol. 12, no. 1, pp. 45–52, 2020.
- [5] M. Yuliana, "Digitalisasi Sistem Inventori pada UMKM: Studi Kasus di Yogyakarta," Jurnal Sistem Informasi Bisnis, vol. 8, no. 1, pp. 31–38, 2021.
- [6] S. Handayani, T. Rachmawati, dan E. Prabowo, "Pentingnya Digitalisasi Inventori pada UMKM," Jurnal Ekonomi dan Kewirausahaan, vol. 10, no. 3, pp. 65–72, 2022.
- [7] I. W. Astuti dan A. Prasetyo, "Peran LSTM dalam Prediksi Permintaan pada UMKM," Jurnal Informatika dan Sains, vol. 6, no. 2, pp. 88–95, 2021.
- [8] H. Fitriani, "Prediksi Jumlah UMKM Menggunakan LSTM," Jurnal Teknologi dan Informasi, vol. 9, no. 1, pp. 22–29, 2022.
- [9] A. Lestari dan F. Zulfikar, "Forecasting Harga Cabai Merah Menggunakan LSTM," Jurnal Agribisnis Digital, vol. 5, no. 2, pp. 101–108, 2021.
- [10] D. Kurniawan dan S. Lestari, "Prediksi Harga Minyak Mentah Dunia Menggunakan LSTM," Jurnal Energi dan Informatika, vol. 7, no. 1, pp. 56–63, 2020.
- [11] R. Kusuma, D. Setiawan, dan E. Wulandari, "Prediksi Stok Produk Sari Roti Menggunakan LSTM," Jurnal Sistem Cerdas, vol. 4, no. 3, pp. 144–151, 2021.
- [12] A. Susilo dan B. Rahardjo, "Optimasi Distribusi Logistik Menggunakan LSTM," Jurnal Logistik dan Teknologi, vol. 9, no. 2, pp. 77–83, 2022.
- [13] N. Wijayanti, "Penerapan Algoritma C4.5 untuk Prediksi Kebutuhan Stok Barang," Jurnal Informatika, vol. 6, no. 2, pp. 94–100, 2020.
- [14] M. Azhari dan F. Haryanto, "Perbandingan LSTM dan Metode Konvensional dalam Prediksi Harga Bawang Putih," Jurnal Teknologi Pertanian, vol. 11, no. 1, pp. 35–42, 2022.
- [15] S. Indrawati, "Keterbatasan Metode Tradisional dalam Prediksi Stok UMKM," Jurnal Riset Manajemen dan Bisnis, vol. 10, no. 1, pp. 17–24, 2021.
- [16] R. Hardiansyah, "Tingkat Adopsi Teknologi AI pada UMKM Skala Mikro," Jurnal Inovasi Digital UMKM, vol. 3, no. 2, pp. 55–62, 2022.
- [17] A. Dewi dan H. Santoso, "Karakteristik dan Penanganan Nira Sebagai Bahan Baku," Jurnal Pangan dan Agroindustri, vol. 8, no. 1, pp. 33–39, 2021.