Implementing LSTM-Based Deep Learning for Forecasting Food Commodity Prices with High Volatility: A Case Study in East Java Province

A I E Nensi[1,*], W Pangesti[1], N Syukri[1], M Al-Maida[1], and K A Notodiputro[1]

[1]Departement of Statistics and Data Science, IPB Univeriesty

*Corresponding author's email: andiillaervianinensi@apps.ipb.ac.id

Abstract. Accurate food price forecasting is essential for maintaining market stability and food security. East Java Province was selected as the study area because it is one of Indonesia's main food production centers and a major contributor to national inflation. This study compares three deep learning architectures LSTM, Bi-LSTM, and hybrid CNN-LSTM to forecast the prices of four key food commodities (red chili, shallots, medium-grade rice, and beef) in East Java. Hyperparameter tuning was performed using grid search, and performance was evaluated using MAPE, MAE, and RMSE. The results show that the Bi-LSTM model consistently provides the best performance compared to LSTM and CNN-LSTM across the four analyzed commodities. Based on MAPE, MAE, and RMSE values, Bi-LSTM achieved the lowest forecasting errors for all commodities. The MAPE values of Bi-LSTM were 1.73% for red chili, 0.60% for shallots, 0.23% for medium-grade rice, and 0.08% for beef, all of which were lower than those of LSTM and CNN-LSTM models. These findings highlight Bi-LSTM's bidirectional architecture, which leverages contextual information from both past and future data sequences, making it the most robust and effective model for forecasting food prices under varying volatility. The study provides practical insights for policymakers and supply chain stakeholders in supporting price stability and food security.
Keyword: Bi-LSTM, CNN-LSTM, Deep Learning, Food Price Forecasting, LSTM

1. Introduction

Food price stability is a fundamental pillar in maintaining Indonesia's food security and economic stability. Fluctuations in the prices of basic commodities contribute significantly to inflation, which has an impact on the decline in people's purchasing power, especially low-income households [1]. This condition not only affects the economic sector but also has an impact on social welfare and public perception of the reliability of daily necessities. Stabilizing food prices aligns with the Sustainable Development Goals (SDGs), particularly SDG 2 (Zero Hunger) and SDG 8 (Decent Work and Economic Growth), which focuses on sustainable and inclusive economic growth.

East Java Province occupies a strategic position as one of the national food barns. In 2024, rice production in East Java reached 9.65 million tons of milled dry grain (GKG) or equivalent to 5.35 million tons of rice, making it a major contributor to the national food supply [2]. In addition to rice, East Java is also a center for the production of important horticultural commodities, such as red chili (562,000 tons) and shallots (484,000 tons). This province is also the main supplier of national animal protein needs through meat and poultry commodities. With such a significant contribution, food price fluctuations in East Java not only affect the local community but also have direct implications for national food stability. These four commodities are consistently recorded by the Central Statistics Agency (BPS) and the Ministry of Agriculture as the main contributors to food inflation [3]. BPS data (June 2025) recorded East Java's inflation at 0.43% (month-to-month), mainly triggered by increases in the prices of red cachili, rice, shallots, and chicken eggs, which shows the crucial role of the food sector in inflation [4].

The frequent fluctuations in food prices (high market volatility) make conventional modeling approaches less effective for predicting price levels, as they often fail to capture nonlinear patterns in time-series data [5]. Given these limitations, deep learning methods have emerged as effective alternatives. Deep learning provides a way which is robust for interpreting and modeling time series data, with better data complexity handling capabilities. One widely used architecture is Long Short-Term Memory (LSTM), which was developed to overcome the vanishing gradient problem in Recurrent Neural Network (RNN) and has proven effective in learning long-term dependencies. Previous studies, particularly in forecasting highly volatile stock prices, have shown that conventional models such as ARIMA tend to yield lower accuracy compared to deep learning models like LSTM when dealing with nonlinear and unstable price dynamics [6]. However, standard LSTM only processes data sequentially forward, so it does not utilize the context of future data. To address this limitation, Bidirectional LSTM (Bi-LSTM) incorporates both forward and backward passes, enabling more comprehensive temporal feature extraction. Similar findings have been reported in other domains, such as such as energy consumption forecasting, Bi-LSTM outperforms LSTM and traditional models like ARIMA and SARIMAX, confirming its superiority in modeling nonlinear time series data [7] . Recent developments have also integrated Convolutional Neural Networks (CNN) with LSTM in a hybrid CNN–LSTM architecture, where CNN extracts local patterns or important features from the data before it is further processed sequentially by LSTM [8].

This study compares three deep learning architectures—LSTM, Bi-LSTM, and CNN-LSTM—to predict prices of four major food commodities in East Java. The objective is to provide data-driven insights that support policymakers in maintaining price stability and protecting consumers from harmful fluctuations.

2. Research Method

2.1. Long Short-Term Memory (LSTM)

LSTM employs a special gate mechanism that controls access to information in memory cells, enabling LSTM to handle and learn long-term historical data patterns [9]. The following figure 1 illustrates the architecture of an LSTM, which consists of an input gate, an output gate, and a forget gate.

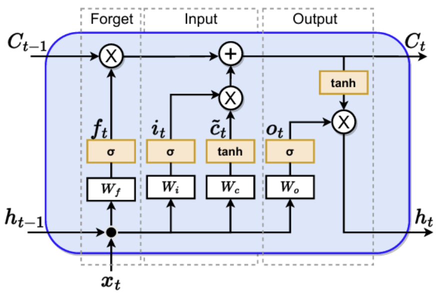


Figure 1. LSTM architecture [10]

$$i_t = \sigma(W_i.[h_{t-1} + x_t] + b_i) \quad (1)$$
$$f_t = \sigma\left(W_f.[h_{t-1} + x_t] + b_f\right) \quad (2)$$
$$o_t = \sigma(W_o.[h_{t-1} + x_t] + b_o) \quad (3)$$
$$\widetilde{C_t} = tanh(W_c.[h_{t-1} + x_t] + b_c) \quad (4)$$
$$C_t = f_t.C_{t-1} + i_t.\widetilde{C_t} \quad (5)$$

The input gate (i_t) determines the new information added to the cell state, and its new candidate value is obtained through the tanh activation function$(\tilde{C}_t)$. These two components are then used to update the cell state (C_t). The forget gate (f_t) functions to decide which information from the previous cell state (C_{t-1}) should be retained or discarded through the sigmoid activation function. Finally, the output gate (O_t) determines which information is output from the cell state, where the final hidden state value is obtained from the product of the output gate and the tanh activation of the cell state. Thus, LSTM is able to retain long-term information while adding new patterns from sequential data.

2.2. Bidirectional Long Short-Term Memory (Bi-LSTM)

Bi-LSTM is a bidirectional combination of LSTM. Bi-LSTM processes data sequences forward (from the beginning to the end of the sequence) and backward (from the end to the beginning of the sequence) [11]. In this way, Bi-LSTM can utilize past and future information simultaneously, resulting in a richer context representation. The Bi-LSTM structure consists of two LSTM layers running in parallel in different directions, which are then combined in the output layer [12]. The following figure 2 illustrates the structure of the Bi-LSTM layer. Both layers have the same output.

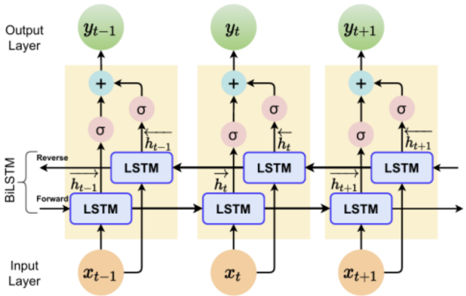


Figure 2. Bi-LSTM architecture [10]

$$\overrightarrow{h_t} = \mathcal{L}\left(x_t W_x \vec{h}_t + \overrightarrow{h_{t-1}} \mathrm{W}_{\vec{h}\vec{h}} + \mathrm{b}_{\vec{h}}\right) \tag{6}$$

$$\overleftarrow{h_t} = \mathcal{L}\left(x_t W_{x\overleftarrow{h}} + \overleftarrow{h_{t-1}} W_{\overleftarrow{h}\overleftarrow{h}} + b_{\overleftarrow{h}}\right) \tag{7}$$

$$y_t = \mathcal{L}\left(\vec{h}_t W_{\vec{h}y} + \overleftarrow{h}_t W_{\overleftarrow{h_t}y} + b_y\right) \tag{8}$$

Essentially, Bi-LSTM uses the same principle as LSTM, namely three gates in measuring information flow. The data processing involves current input (x_t), hidden state from the previous time (h_{t-1}), activation function $(\mathcal{L})$, weights (W) and bias (b). The difference lies in the direction of propagation, where there are two processing paths. The forward hidden path $\left(\vec{h_t}\right)$ utilizes information from the previous hidden state to process the input at the current step. Conversely, the backward hidden path $\left(\overleftarrow{h_t}\right)$ processes the input by considering information from the hidden state at the next step. The results of these two paths are then combined to produce the final output (y_t) at each time step. Thus, Bi-LSTM is able to capture both past and future contexts in a data sequence.

2.3. Convolutional Neural Network (CNN-LSTM)

CNN-LSTM is a hybrid architecture that combines the advantages of Convolutional Neural Network (CNN) and Long Short-Term Memory (LSTM) [13]. CNN is used to extract local features from sequential data, while LSTM utilizes these feature representations to learn long-term temporal dependencies. With this combination, CNN-LSTM is effective in capturing complex patterns of both short-term and long-term dependencies [14]. CNN works by using one-dimensional convolution (1D convolution) to capture local patterns or short-term dependencies, such as trends or cycles in time series data [15]. This combination makes CNN-LSTM effective in identifying complex patterns, both short-term dependencies and long-term dependencies, so it is often used in prediction and anomaly detection tasks on time series data. The following figure 3 shows the architecture of CNN-LSTM.

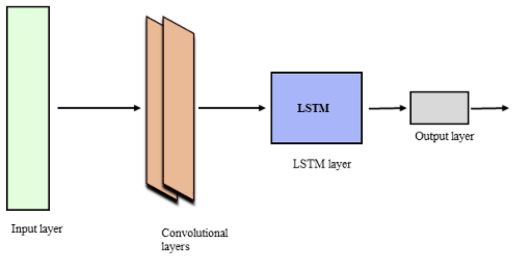


Figure 3. CNN-LSTM architecture [16]

In general, the CNN-LSTM processing stages are CNN processing, where sequential data is fed into the CNN to generate feature maps that extract local patterns from each time step. Next is sequence processing, where the feature maps obtained are then fed into the LSTM, which learns temporal dependencies and updates the hidden state at each time step to capture long-term dependencies. The output from the LSTM is used to generate time series predictions.

2.4. Data

The primary focus of this study is on commodity prices, specifically the prices of red chili, shallots, and beef in East Java Province, from 2014 to 2024, presented as daily time-series data. The data were obtained from the Information System on Availability and Price Development of Basic Materials in East Java (SISKAPERBAPO) website, through a scraping process [17]. The dataset is divided into training data (80%) and testing data (20%). The training set is used to train the forecasting models and capture temporal patterns, while the testing set is used to evaluate the models' generalization ability and assess predictive performance on previously unseen data. These commodities were selected based on data from Statistics Indonesia (BPS) and the Ministry of Agriculture, which show that red chili, shallots, rice, and beef are the main contributors to food inflation in East Java. Their strong price fluctuations make them relevant for analyzing volatile food price patterns.

2.5. Analysis Methodology

The analysis was conducted using LSTM, Bi-LSTM, and CNN-LSTM models separately but with the same procedure:

1. Exploring the data to see an overview of the characteristics of each commodity price.
2. Performing data transformation using the min-max scaler method. This is done to equalize the data scale in the range of $0 - 1$ without changing the original data distribution.
3. Converting the data into three dimensions: number of samples (n), time step as the prediction time window prediction, and the number of variables in each sample. Modeling with LSTM, Bi-LSTM, and CNN-LSTM
   a. Dividing the training data and test data
   b. Initialize hyperparameters for each model as shown in table 1.

Table 1. LSTM-Based models parameters.

Hyperparameter	Value
Number of layers	1
Number of neurons	[32, 50, 100]
Batch size	[16, 32]
Activation Function	Tanh, ReLU
Learning rate	[0.01, 0.001]
Epochs	100

   c. Performing hyperparameter tuning to obtain the best hyperparameter combination. Hyperparameter tuning is performed using grid search.
   d. Calculating model evaluation for each hyperparameter combination and selecting the best hyperparameter combination based on the RMSE loss function
   e. Performing modeling using the best hyperparameters for each model.

f. Predicting each commodity price with each model, resulting in $4 \times 3 = 12$ models.

4. Denormalization to restore the previous values from the range 0− 1 to the values the actual range
5. Evaluate the performance of the three models using the MAPE, MAE, and RMSE metrics, with the best model selected based on the smallest value.

Figure 4 below is a flowchart of the analysis.

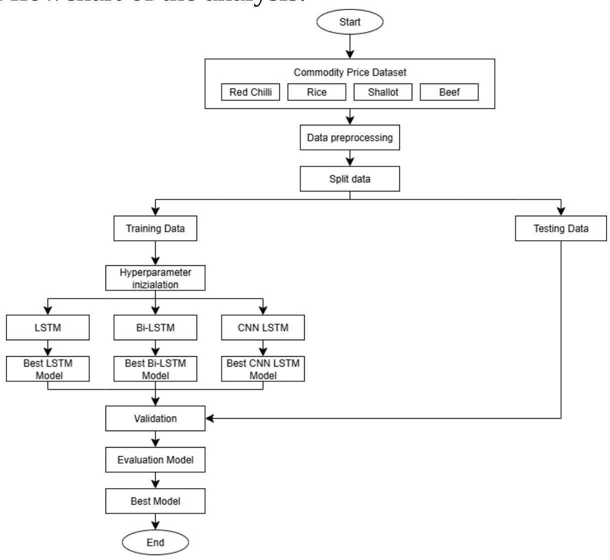
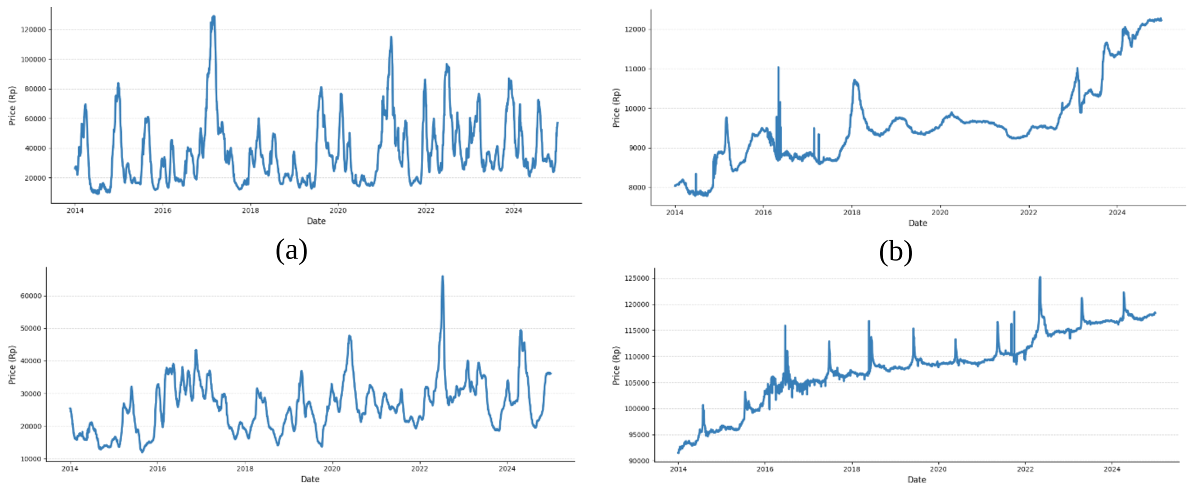


Figure 4. Research stage

3. Result and Discussion

3.1. Descriptive Analysis

Descriptive analysis was conducted to provide an overview of the patterns, trends, and variations in food commodity prices in East Java. This analysis aimed to identify differences in the level of fluctuation between commodities and assess price stability during the observation period.



(a) (b)

(c) (d)

Figure 5. Daily price trends of (a) red chili, (b) medium rice, (c) shallot, and (d) beef in East Java, 2014–2025.

Figure 5 shows the daily price trends for four food commodities in East Java for the period 2014–2025. The price of red chili showed the sharpest fluctuations, with a peak price of 129,068 rupiah/kg on February 26, 2017, and a low of 9,009 rupiah/kg on July 7, 2014. These fluctuations can occur because red chili is highly sensitive to climatic conditions and pest infestations, while its harvesting period is relatively short. In addition, its perishable nature and short shelf life mean that even small changes in supply can trigger drastic price spikes. Shallots also experienced significant fluctuations, marked by a price spike of up to 65,994 rupiah/kg on July 13, 2022, and a minimum price of 11,889 rupiah/kg on August 26, 2015. Meanwhile, beef prices are relatively stable with a slow upward trend, with the lowest price of 91,464 rupiah/kg on January 3, 2014, reaching a peak of 125,228 rupiah/kg on May 6, 2022. These differences become more evident when examined through the descriptive statistics presented in table 2.

Table 2. Descriptive statistics of commodity prices

Commodities	Mean	std	Min	25%	50%	75%	Max
Red chili	37099.79	22011.51	9009	19611.2	31767.5	48055.5	129068
Shallot	26093.72	7985.63	11889	19853	25864	31041.25	65994
Medium rice	9591.73	1034.99	7777	8862	9482.5	9784	12273
Beef	107978	6892.14	91464	105021	108478	114037.75	125228

Table 2 provides an important overview of the price movements of four commodities. The price of red chili has an average of 37099.79 with a standard deviation of 22011.51 and a wide price range, from 9009 to 129068 rupiah/kg, which confirms high volatility. Shallots also exhibited considerable fluctuations, with an average of 26,093.72 and a standard deviation of 7,985.63, as well as a fairly wide price range between 11,889 and 65,994 rupiah/kg. In contrast, medium-grade rice was relatively stable, with an average of 9591.73, a low standard deviation of 1034.99, and a narrow price range of 7777–12273 rupiah/kg, indicating that price movements were relatively controlled. Beef was the most consistent commodity, with the highest average price of 107,978 but a small standard deviation of 6,892.14 and a relatively narrow price range of 91,464–125,228 rupiah/kg.

Overall, both visual patterns and descriptive statistics show that red chili is the commodity with the highest price volatility, followed by shallots. Conversely, medium-grade rice and beef are relatively stable throughout the observation period.

3.2. Prediction Analysis of LSTM, Bi-LSTM, and CNN-LSTM

This study compares the performance of three deep learning architectures —namely, LSTM, Bi-LSTM, and CNN-LSTM —in predicting the prices of strategic food commodities. The comparison was conducted to identify the model that is most capable of capturing price movement patterns in both highly volatile and relatively stable data.

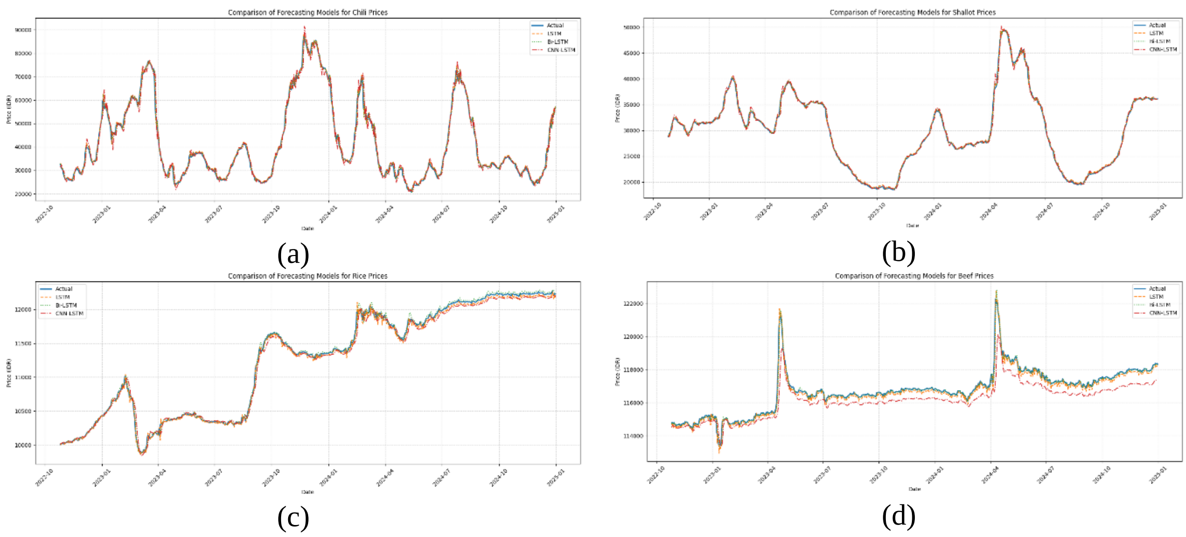


Figure 6. Comparison of the best LSTM, Bi-LSTM, and CNN-LSTM models for (a) red chili, (b) shallot, (c) medium rice, and (d) beef.

Figure 6 presents a comparison of price forecasts for red chili, shallots, medium-grade rice, and beef using three deep learning architectures. In general, all three models are able to capture price movement patterns effectively and accurately follow actual trends based on historical data. However, there are variations in performance between models, especially for commodities with high price volatility.

Volatile commodities, such as red chili (Figure 6a) and shallots (Figure 6b), exhibit more pronounced differences between models. The fluctuating price of red chili indicates that Bi-LSTM predictions are smoother and closer to actual values than those of LSTM and CNN-LSTM. A similar pattern is also observed in shallots, where Bi-LSTM produces a more consistent prediction line following the actual data, while CNN-LSTM appears to have a greater deviation during periods of price spikes, indicating its inability to capture these anomalies. The standard LSTM occupies a position between the two.

For more stable commodities, such as medium-grade rice (Figure 6c) and beef (Figure 6d), all three models show similar accuracy. However, Bi-LSTM remains superior with more precise predictions, especially during the rise in rice prices at the end of 2024 and the surge in beef prices in mid-2023. In contrast, CNN-LSTM for beef shows an underestimate bias, as its convolutional architecture tends to filter sporadic price spikes as noise. In fact, these spikes are often triggered by non-cyclical external factors, such as holidays, disease outbreaks, or import policies. As a result, CNN-LSTM only captures stable underlying patterns and fails to represent price spikes. These findings confirm that hybrid architectures are not always the best solution for all types of time series data.

In contrast, the advantage of Bi-LSTM comes from its bidirectional mechanism. This architecture not only processes information from the past to the future (forward) but also from the future to the past (backward). This dual perspective provides richer temporal context for each data point, enabling the model to account for influences from both preceding and subsequent trends. This ability makes Bi-LSTM particularly robust in capturing complex temporal dynamics, both subtle fluctuations and drastic changes, resulting in more accurate representations than unidirectional models such as LSTM and hybrid CNN-LSTM.

3.3. Model Evaluation Comparison

The models were evaluated by comparing their performance in forecasting the prices of four food commodities in East Java. The performance of each model was evaluated using three metrics: Mean Absolute Percentage Error (MAPE), Mean Absolute Error (MAE), and Root Mean Squared Error (RMSE). A summary of the evaluation results is shown in table 3.

Table 3. Prediction performance metrics of LSTM, Bi-LSTM, and CNN-LSTM for each commodity.

	Model	Metrics Evaluation		
		MAPE (%)	MAE	RMSE
Red Chili	LSTM	2.06	830.22	1189.22
	Bi-LSTM	**1.73**	**724.27**	**1133.20**
	CNN-LSTM	3.69	1542.07	2285.53
Shallot	LSTM	0.79	226.46	321.95
	Bi-LSTM	**0.60**	**184.62**	**296.47**
	CNN-LSTM	1.09	339.39	521.07
Medium Rice	LSTM	0.25	28.84	38.57
	Bi-LSTM	**0.23**	**25.48**	**35.98**
	CNN-LSTM	0.39	44.60	60.88
Beef	LSTM	0.13	153.59	207.52
	Bi-LSTM	**0.08**	**88.88**	**183.14**
	CNN-LSTM	0.50	591.51	736.13

The evaluation results in table 3 show that the Bi-LSTM model consistently outperforms the other models for almost all commodities by generating the lowest MAPE, MAE, and RMSE values. The Bi-LSTM's bidirectional architecture enables it to capture both past and future price patterns, leading to superior predictive accuracy. For instance, when predicting the highly volatile red chili prices, Bi-LSTM achieved a MAPE of 1.73%, compared to the standard LSTM (2.06%) and CNN-LSTM (3.69%). For shallots, Bi-LSTM again performed best, with the lowest MAPE of 0.60%.

For commodities with relatively stable price patterns, such as medium-grade rice and beef, all models performed well, with MAPE values below 1%. Notably, Bi-LSTM again proved to be the most accurate, achieving MAPE values of 0.23% for rice and 0.08% for beef, indicating its strength in handling stable price data and capturing temporal dependencies. Meanwhile, the CNN-LSTM hybrid model demonstrated lower performance across all commodities, recording the highest prediction error; this was especially evident for red chili (MAPE: 3.69%), suggesting its limitations in modeling such patterns.

These findings indicate that the bidirectional architecture of Bi-LSTM, which utilizes contextual information from both directions (forward and backward) in data sequences, is more effective in capturing complex temporal patterns in price data. This finding aligns with research conducted by [14], which suggests that Bi-LSTM is superior to other time series methods in predicting fresh vegetable sales. Even when integrated with price strategy and restocking optimization using Particle Swarm Optimization (PSO), it is able to provide maximum profit. This effectiveness applies to both highly volatile and stable commodities. The results of this study extend the advantages of Bi-LSTM to food price prediction, demonstrating its effectiveness on both highly volatile and stable data. Conversely, the relatively weaker performance of CNN–LSTM demonstrates that hybrid architectures are not inherently superior, and model selection should be adapted to the specific characteristics of the data.

3.4. *Best Model Prediction Results*

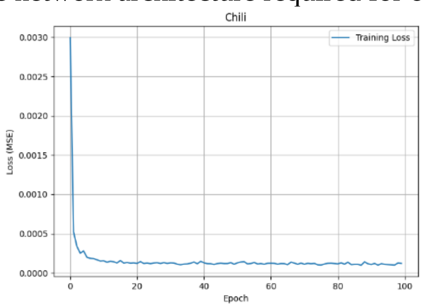
In this study, the Bi-LSTM hyperparameter tuning process was carried out using the grid search method to identify the optimal configuration for forecasting the prices of cayenne pepper, shallots, rice, and beef. The grid search process was conducted based on the hyperparameter combinations listed in table 1. The final results of this process are presented in table 2, which lists the optimal hyperparameter combinations from the Bi-LSTM model for each commodity.

Table 4. The best hyperparameter combination of the BI-LSTM model

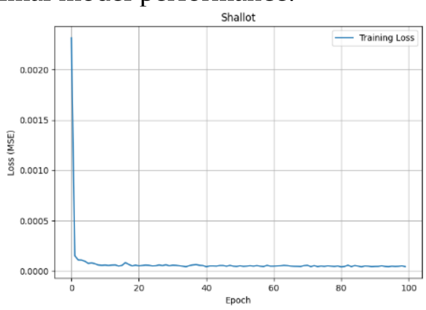
Commodities	Number of layers	Number of neurons	Batch size	Activation Function	Learning rate
Red Chili	1	50	16	Tanh	0.01
Shallot	1	50	16	Tanh	0.01
Medium Rice	1	50	16	Relu	0.01
Beef	1	32	32	Relu	0.01

The tuning results show that the optimal configuration is relatively consistent across most commodities, although there are certain variations. For cayenne pepper, shallots, and medium-grain rice, the optimal combination consists of one layer with 50 neurons and a batch size of 16, using the Tanh activation function for cayenne pepper and shallots, and the ReLU function for rice.

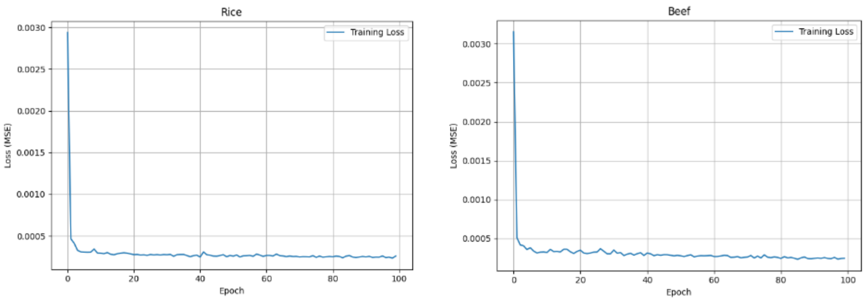
Meanwhile, beef uses a different setup: 32 neurons, a batch size of 32, and a ReLU activation function. All commodities share the same learning rate of 0.01. These configuration differences demonstrate that the data characteristics for each commodity, such as price volatility and stability, influence the network architecture required for optimal model performance.



(a)



(b)



(c) (d)

Figure 7. Training loss curves of the Bi-LSTM for (a) red chili, (b) shallot, (c) medium rice, and (d) beef prices.

Figure 7 shows the training loss curve of the Bi-LSTM model for the four food commodities during 100 training epochs. For all commodities, the training loss value decreased dramatically in the first ten epochs before finally stabilizing at a very low value close to zero. This demonstrates the model's value in efficiently capturing and learning price patterns in commodity data. Overall, these results confirm that the selected hyperparameter configuration produces a model that converges quickly without any indication of overfitting, highlighting Bi-LSTM's stability and effectiveness in modeling each commodity's price patterns.

4. Conclusion

This study addresses the critical need for accurate food price forecasting by comparing deep learning models. The results show that Bi-LSTM consistently achieved the best performance with the lowest error across all four commodities. Bi-LSTM's advantage lies in its ability to process temporal information from both past and future price points, enabling it to understand price dynamics in a greater context and identify complex patterns in both volatile and stable markets. This bidirectional approach enhances its accuracy over standard LSTM, which only considers past values, and over CNN-LSTM, which focuses more on spatial feature extraction. In contrast, CNN-LSTM showed suboptimal performance, particularly with a tendency to underestimate spikes in beef prices. These findings emphasize that more complex architectures do not always guarantee better results, and model selection should be adapted to the characteristics of the time-series data. Practically, Bi-LSTM has the potential to serve as the basis for early warning systems and decision support in maintaining price stability and food security. In the future, further research should focus on integrating external variables, such as macroeconomic and climate factors, to enhance the predictive capabilities of Bi-LSTM.

Acknowledgement

The authors gratefully acknowledge the financial support provided by the Indonesia Endowment Fund for Education (LPDP), Ministry of Finance of the Republic of Indonesia.

References

[1] M. R. Widarso and S. Djamaluddin, "ANALISIS HARGA PANGAN POKOK TERHADAP KETAHANAN PANGAN DI INDONESIA," *SEPA: Jurnal Sosial Ekonomi Pertanian dan Agribisnis*, vol. 21, no. 2, p. 256, Oct. 2024, doi: 10.20961/sepa.v21i2.83821.

[2] BPS Provinsi Jawa Timur, "ringkasan-eksekutif-luas-panen-dan-produksi-padi-di-jawa-timur-2024," vol. 4, Jawa Timur: BPS Provinsi Jawa Timur, 2025.

[3] Yeniartha, "Panen Raya Serentak, Jatim Siap Siaga Sebagai Lumbung Pangan Nasional," Balai Besar Pelatihan Pertanian. Accessed: Aug. 29, 2025. [Online]. Available: https://bbppketindan.bppsdmp.pertanian.go.id/blog/post/panen-raya-serentak-jatim-siap-siaga-sebagai-lumbung-pangan-nasional

[4] Badan Pusat Statistik Kabupaten Probolinggo, "Perkembangan Indeks Harga Konsumen Provinsi Jawa Timur Desember 2024," *Badan Pusat Statistik*, no. 32, 2025, [Online]. Available:

https://probolinggokab.bps.go.id/id/pressrelease/2025/01/08/255/perkembangan-indeks-harga-konsumen-provinsi-jawa-timur-desember-2024.html

[5] W. Pangesti, N. Syukri, K. A. Notodiputro, Y. Angraini, L. Nissa, and A. Mualifah, “Performance Evaluation of ARIMA,” 2025.

[6] A. Illa, E. Nensi, M. Al Maida, K. A. Notodiputro, Y. Angraini, and L. Nissa, “Performance Analysis of ARIMA, LSTM, and Hybrid ARIMA-LSTM in Forecasting the Composite Stock Price Index,” vol. 10, no. 2, pp. 588–604, 2025.

[7] H. Alizadegan, B. Rashidi Malki, A. Radmehr, H. Karimi, and M. A. Ilani, “Comparative study of long short-term memory (LSTM), bidirectional LSTM, and traditional machine learning approaches for energy consumption prediction,” *Energy Exploration and Exploitation*, vol. 43, no. 1, pp. 281–301, Jan. 2025, doi: 10.1177/01445987241269496.

[8] W. Bao, W. Su, X. Zhao, and J. Zhuang, “A study on short-term vegetable price prediction based on the CNN-LSTM-Attention model,” *Discover Food*, vol. 5, no. 1, 2025, doi: 10.1007/s44187-025-00426-2.

[9] N. Bacanin *et al.*, “Multivariate energy forecasting via metaheuristic tuned long-short term memory and gated recurrent unit neural networks,” *Inf Sci (N Y)*, vol. 642, Sep. 2023, doi: 10.1016/j.ins.2023.119122.

[10] B. Bohara, R. I. Fernandez, V. Gollapudi, and X. Li, “Short-Term Aggregated Residential Load Forecasting using BiLSTM and CNN-BiLSTM,” in *2022 International Conference on Innovation and Intelligence for Informatics, Computing, and Technologies, 3ICT 2022*, Institute of Electrical and Electronics Engineers Inc., 2022, pp. 37–43. doi: 10.1109/3ICT56508.2022.9990696.

[11] M. Schuster and K.K. Paliwal, “Bidirectional recurrent neural networks,” *IEEE Transactions on Signal Processing*, vol. 45, no. 11, p. 26732681, 1997.

[12] A. Zeroual, F. Harrou, A. Dairi, and Y. Sun, “Deep learning methods for forecasting COVID-19 time-Series data: A Comparative study,” *Chaos Solitons Fractals*, vol. 140, Nov. 2020, doi: 10.1016/j.chaos.2020.110121.

[13] W. Lu, J. Li, Y. Li, A. Sun, and J. Wang, “A CNN-LSTM-based model to forecast stock prices,” *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/6622927.

[14] L. Zhen and A. Bărbulescu, “Comparative Analysis of Convolutional Neural Network-Long Short-Term Memory, Sparrow Search Algorithm-Backpropagation Neural Network, and Particle Swarm Optimization-Extreme Learning Machine Models for the Water Discharge of the Buzău River, Romania,” *Water (Switzerland)*, vol. 16, no. 2, Jan. 2024, doi: 10.3390/w16020289.

[15] P. Lara-Benítez, M. Carranza-García, and J. C. Riquelme, “An Experimental Review on Deep Learning Architectures for Time Series Forecasting,” *Int J Neural Syst*, vol. 31, no. 3, Mar. 2021, doi: 10.1142/S0129065721300011.

[16] R. Murugesan, E. Mishra, and A. H. Krishnan, “Deep Learning Based Models: Basic LSTM, Bi LSTM, Stacked LSTM, CNN LSTM and Conv LSTM to Forecast Agricultural Commodities Prices,” Nov. 01, 2021. doi: 10.21203/rs.3.rs-740568/v1.

[17] Disperindag Jatim, “Sistem Informasi Ketersediaan dan Perkembangan Harga Bahan Pokok di Jawa Timur,” Siskaperbapo.