International Journal Software Engineering and Computer Science (IJSECS)
5 (1), 2025, 215-225
Published Online April 2025 in IJSECS (http://www.journal.lembagakita.org/index.php/ijsecs)
P-ISSN: 2776-4869, E-ISSN: 2776-3242. DOI: https://doi.org/10.35870/ijsecs.v5i1.3526.

RESEARCH ARTICLE Open Access

Forecasting Chili Prices in Metro City Using Long Short-Term Memory (LSTM)

Gusti Made Gunadi *
Informatics Engineering Study Program, Faculty of Business Technology and Science, Universitas Dharma Wacana, Metro City, Lampung Province, Indonesia.
Corresponding Email: kadekgunadi1212@gmail.com.

Andreas Perdana
Informatics Engineering Study Program, Faculty of Business Technology and Science, Universitas Dharma Wacana, Metro City, Lampung Province, Indonesia.
Email: andreas.perdana@dharmawacana.ac.id.

Received: December 24, 2024; Accepted: February 1, 2025; Published: April 1, 2025.

Abstract: Cayenne pepper is one of the important commodities in the staple market in Indonesia which has a vital role in people's daily lives. Fluctuations in the price of cayenne pepper are often a challenge that impacts farmers and consumers, causing uncertainty in production and distribution planning. This research aims to develop a cayenne pepper price prediction model using the Long Short-Term Memory (LSTM) method, utilizing historical data from the data.metrokota.go.id portal for the period October 2023 to October 2024. By using LSTM, this model successfully captures long-term patterns in cayenne price data, with a Final Validation Loss of 0.00249 which indicates a high level of accuracy. The prediction results are expected to help farmers determine the optimal selling time, traders in managing stocks efficiently, and policy makers in formulating strategies to mitigate the impact of price fluctuations. In addition, this study highlights practical implications for stabilizing commodity markets, particularly in Metro City, as well as the relevance of these findings to be applied to other agricultural commodities.

Keywords: Cayenne Pepper; Price Prediction; Long Short-Term Memory; Historical Data; Metro City.

1. Introduction

One of the commodity strategies in staple food market in Indonesia is Chili pepper. It is an important ingredient of various traditional recipes and its presence in daily consumptions of common people. Having a high demand from both the household sector, culinary businesses and food processing industry, chili pepper is one of the major components in the national food supply chain. Yet farmer as main producer and end users the most farmer faced with big challenge for any way it is frequent price instability. There is another effect, the any move on the price of chili pepper affects not only purchasing power of population but also have a huge effect on national inflation rate. These causes are countless and include production dynamics, distribution bottlenecks, seasonal or economic shifts in market demand for interacting with the market. Moreover, external reasons like the existence of extreme weather and transport problems often make the situation worse thus making price forecasting even more complicated [1]. This ambiguity makes it hard for farmers to design optimal production strategies, while traders may also be at a risk of loss with improper stock management. Commodity prices are unstable within this range as well so there is a glut on macroeconomic policy. Because accurate forecast data is not available, market stabilization in different manners based such as price control or national stock control practice is necessary by government sometimes these mechanisms are not completely effective. This indicates an actual need for some kind of systematic study of price trends and forecasting future changes. Ahmmedet *et al.*, (2023) A historical data based approach might be a way to reduce the risks due to price volatility, and provide a foundation for well – informed decisions for each stake-holders [2].

There has been an overall surge in the application of artificial intelligence—focused data analysis methods through machine learning techniques in an attempt to anticipate patterns of Agricultural commodity prices, as the means to overcome these challenges. One of the most recognizable is the Long Short-Term Memory (LSTM), a recurrent neural network designed to model long-term dependencies in sequences. Compared to traditional methods (ARIMA or linear regression) LSTM can perform better at recognizing non-linear patterns and complex relationships among historical data in commodities, leading to more accurate predictions for the commodities which have higher price volatility [3][4]. Past researches indicate that LSTM can outperform the conventional methods to forecast food commodity prices with better performance especially in high varied data [5]. This includes also the local conditions of cayenne pepper market in Indonesia (seasonal and geographic). Cayenne pepper production is heavily influenced by the Central and South American tropical climate in which it grows in different regions with variable planting and harvest patterns. Inequality in regional distribution is another of the most significant causes for price instability, especially in urban areas such as Metro City, which is the main scope of this study. The data in of the data official portal data.metrokota.go.id proves that cayenne pepper prices have been fluctuating in a certain range over time, which helps to implement broader challenges of handling agricultural commodities on local level [6]. In this research, the objective is to develop a cayenne pepper price prediction model using LSTM method based on data obtained from data.metrokota.go.id portal of historical data. The newly created model is anticipated to be useful tool for farmers on optimal time selection of planting and harvesting, traders for better management of their stock as well as policy maker for the formulation of more precise stabilization price measure. Consequently, this effort is expected to contribute measurable economic benefits by supporting the building of stability for the cayenne pepper commodities market. Also, the outputs of this study could be a reference for the development of the other price prediction model in other agricultural commodities since there are some similar challenges for the agricultural sector in general in Indonesia [7]. Nevertheless, LSTM based approach also comes with its own set of challenges. Model requires enough and long time historical data to make proper predictions. The data deficiency and limitation in data quality can limit the ability of the model not being able to learn long term patterns. Accordingly, this study will also measure how much of these data to apply to LSTM and point toward solutions for possible improvements that help in the prediction accuracy achievement of future forecasts (Li *et al.*, 2023). This research uses critical review of the strengths and weaknesses of the model to present viable solutions for this dynamic agricultural commodity price.

2. Related Work

Agricultural commodity price prediction has become a major focus in various studies due to its significant impact on economic stability and social welfare. Various analytical approaches have been applied to understand and forecast price patterns that are often unpredictable. One of the traditional methods that is widely used is the AutoRegressive Integrated Moving Average (ARIMA), which is known for its ability to identify linear patterns in time series data. However, the main weakness of ARIMA lies in its inability to capture non-linear relationships that often appear in commodity price data, so that its prediction results are often less accurate when faced with complex fluctuations [8][9]. As an alternative, machine learning-based methods such as Support Vector Regression (SVR) and Artificial Neural Networks (ANN) have shown better performance in analyzing data with non-linear characteristics. SVR excels in modeling non-linear relationships through the use of kernel functions, while ANN is able to recognize more complex patterns through adaptive artificial neural network structures. However, both methods still face challenges in capturing long-term dependencies in time series data, especially when the data has high variability or complex seasonal patterns [2]. These limitations often result in inconsistent predictions, especially for agricultural commodities that are affected by various external factors such as weather and logistics disruptions.

To overcome these shortcomings, recent research has turned to recurrent neural network (RNN)-based approaches, specifically Long Short-Term Memory (LSTM). LSTM is specifically designed to handle long-term dependencies in time series data with a memory mechanism that allows important information to be stored over a longer period of time. This advantage makes LSTM more effective than traditional methods such as ARIMA or even standard ANN in the context of commodity price prediction [3]. With the ability to recognize complex and non-linear patterns, LSTM has been shown to provide more accurate results, especially for data with high volatility such as food prices [4]. Several previous studies have confirmed the superiority of LSTM over conventional approaches. Adi and Sudianto (2020) compared ARIMA and LSTM models in predicting food commodity prices, and found that LSTM was significantly superior in capturing long-term patterns that were not detected by ARIMA [5]. Similarly, Hidayat and Wibisono (2024) applied LSTM to forecast rice prices in Indonesia, with results showing higher accuracy compared to traditional methods, especially in the face of seasonal fluctuations. Another study by Jiang and Liu (2023) highlighted the application of deep learning approaches such as LSTM in agricultural price prediction [1], with an emphasis on the ability of this model to integrate various external variables that affect prices.

Additionally, research by Gao and Chai (2023) explored the use of LSTM-based hybrid networks for commodity price forecasting, combining elements from other models to improve prediction accuracy [6]. This approach demonstrated that LSTMs are not only effective as standalone models, but can also be combined with other techniques to address specific challenges in time-series data. Meanwhile, Wang and Chen (2023) in their systematic review of deep learning for agricultural price prediction, confirmed that LSTMs consistently outperform traditional methods, especially in the context of data with high variability and significant uncertainty [7]. Although LSTMs offer many advantages, they are not without challenges. One major drawback is their reliance on the availability of sufficient and high-quality historical data. Limited or incomplete data can reduce the model's ability to produce accurate predictions, especially in capturing long-term seasonal patterns [10]. Furthermore, the high computational complexity of LSTMs can also be a barrier, especially in the context of large-scale implementations or in regions with limited technological resources [11]. Therefore, it is important to consider data quality and computational requirements when applying this approach in commodity price prediction.

This study also considers recent developments in the field of time series prediction, such as a hybrid model that combines LSTM with an attention mechanism. This approach, as explained by Liu and Zhang (2024), allows the model to focus more on the most relevant parts of the data, thereby improving prediction accuracy in situations with high data noise [12]. In addition, Zhou *et al.* (2024) introduced the Informer model, a Transformer-based approach designed for long time series forecasting, which can be an alternative or complement to LSTM in future research [13]. However, for the context of this study, LSTM remains the primary choice due to its methodological maturity and strong empirical support in the literature on agricultural price prediction [14]. Based on the literature review above, it is clear that the LSTM-based approach offers a more robust solution than traditional methods in forecasting commodity prices. Therefore, this study aims to build a chili price prediction model using the LSTM method, utilizing historical data from the data.metrokota.go.id portal. The prediction results are expected to be a significant tool for farmers in determining production strategies, for traders in managing stocks efficiently, and for policy makers in designing more targeted market interventions. Thus, this effort is expected to be able to support the stability of cayenne pepper prices in the market, which ultimately brings economic benefits to all parties involved [15]. Furthermore, this study also has the potential to be a basis for developing prediction models for other agricultural commodities in Indonesia, given the similar challenges faced by this sector as a whole.

3. Research Method

This study adopts a quantitative approach with the aim of examining the relationship between the variables that affect the price of cayenne pepper. The method used includes the process of collecting historical data, pre-processing the data, building a model, and evaluating the performance of the model. The processed data is then analyzed using the Long Short-Term Memory (LSTM) method to predict price patterns based on past trends. The model is built with several important parameters such as the number of LSTM units, dropout, Adam optimizer, and Mean Squared Error (MSE) loss function. Each step is optimized to ensure the prediction results have a high level of accuracy, so that it can help users in strategic decision making. The parameters chosen were based on previous research and preliminary experiments to match the data characteristics. The number of LSTM units was determined to ensure sufficient model capacity in capturing data patterns, while dropout was applied to reduce the risk of overfitting. The Adam optimizer was chosen for its ability to accelerate model convergence with high stability. In addition, this study also validates the advantages of the LSTM model by comparing its performance against traditional methods such as ARIMA and linear regression. This comparison aims to demonstrate the effectiveness of LSTM in recognizing complex and long-term patterns compared to simpler approaches. The evaluation is done using the Root Mean Squared Error (RMSE) metric to quantitatively measure the prediction error rate, thus ensuring that the selected model has the best performance.

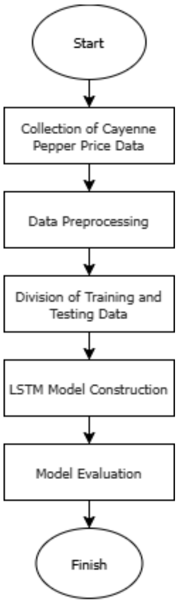

Figure 1. Research Method

3.1 Data Preparation

The data used was obtained from the data.metrokota.go.id portal in the format of daily historical cayenne pepper prices. This data is processed through the stages of cleaning, normalization using MinMaxScaler, and sequence formation for model input.

Table 1. Sample Dataset of Cayenne Pepper Price

No	Item Type	Unit	Price Yesterday (Rp)	Price Today (Rp)	Description	Date
1	Cayenne Pepper	Kg	38.0	38.0	Stable	2023-10-08
2	Cayenne Pepper	Kg	38.0	38.0	Stable	2023-10-09
3	Cayenne Pepper	Kg	38.0	38.0	Stable	2023-10-10
4	Cayenne Pepper	Kg	38.0	38.0	Stable	2023-10-11
5	Cayenne Pepper	Kg	38.0	38.5	Increase	2023-10-12

Once the data is collected, the first step is to perform data cleansing to remove missing or invalid values. Next, the data is normalized using the MinMaxScaler method, which converts the data values to a range of 0 to 1. Normalization is important to remove any bias that may arise due to the different scales of values.

The normalization process is done using the following equation:

$$x' = \frac{x - x_{min}}{x_{max} - x_{min}}$$

Description:
x' : Normalized value
x : Original value
x_{min} : Minimum value in the dataset
x_{max} : Maximum value in the dataset

After the data is normalized, the next step is to form the data into sequences for LSTM model input. The data is then divided into training (80%) and testing (20%) sections. Each sequence consists of a certain amount of historical data that is used to forecast future values.

3.2 LSTM Modeling

Long Short-Term Memory (LSTM) is a type of Recurrent Neural Network (RNN) designed to handle sequential data with the ability to store information in both the short and long term. LSTM overcomes the vanishing gradient problem often experienced by conventional RNNs, making it more reliable in recognizing and predicting patterns in time series data.

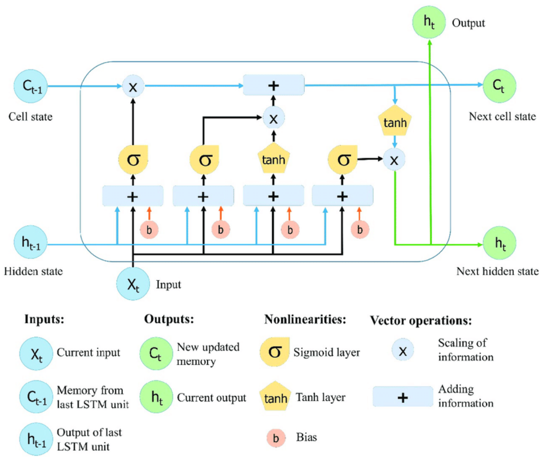

Figure 2. LSTM Unit Structure

LSTM (Long Short-Term Memory) has a main structure consisting of memory cells and several gates that manage the flow of information. The memory cell acts as a long-term store of information, allowing the model to maintain the context of the data over time. This information is managed through three types of gates: input gates, forget gates, and output gates. The input gate serves to regulate the amount of new information to be entered into the memory cells. The information received is processed through the activation function to ensure only relevant information is stored. Meanwhile, the forget gate determines the part of information in the memory that is no longer relevant and needs to be deleted. This process is important to maintain the efficiency and accuracy of the model, especially in handling long series data. Furthermore, the output gate plays a role in organizing the information that is output from the memory cell, which is then used for calculations in the next step. With this mechanism, the LSTM can capture complex patterns and long-term dependencies in sequential data. The combination of these three gates ensures that the model can retain important information, forget irrelevant ones, and produce accurate outputs.

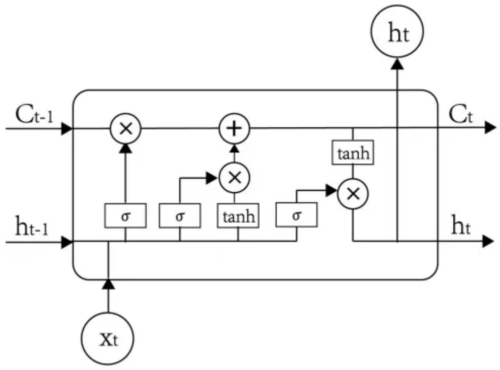

Figure 3. LSTM Unit Diagram

The following are the formulas used in LSTM:

1) Forget Gate

Forget Gate serves to control which information will be forgotten from the previous memory state. In this context, Forget Gate is very important because not all information that has been stored in the previous memory is relevant for the current time step. By using a sigmoid activation function, Forget Gate generates a value between 0 and 1. This value serves to control the extent to which information in the memory cell will be retained or deleted, where a value close to 0 means that the information will be forgotten, while a value close to 1 indicates that the information will be fully retained. The formula for Forget Gate is stated as follows:

$$f_t = \sigma\,(\,W_f \cdot [\,h_{t-1}, x_t\,] + b_f\,)$$

Description:
f_t : the output of Forget Gate at time t,
W_f : Forget Gate weight matrix,
h_{t-1} : the output of the previous hidden state,
x_t :input at time t ,
b_f :bias Forget Gate.

2) Input Gate

Input Gate serves to control which information will be entered into the memory state. This gate is very important because not all new information received at t time needs to be stored. Using a sigmoid activation function, the Input Gate generates a value between 0 and 1 that governs how much new information will be added to the memory. A higher value indicates that more new information will be added, while a lower value means less new information will be added. The formula for Input Gate is expressed as follows:

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

Description:
i_t : the output of the Input Gate at time t ,
W_i : Input Gate weight matrix,
b_i : Gate Input bias.

3) Update Cell State

The Cell State at the time t is updated by considering the forgotten information and the new information to be stored. This process is very important to ensure that the model can keep relevant information and delete unnecessary information. In this way, LSTM can retain important information in the long run. The formula for updating the Cell State is expressed as follows:

$$C_t = f_t \cdot C_{t-1} + i_t \cdot C_t$$

Description:
C_t : Cell State updated at time t
$C_{t(-1}$: Cell State.

4) Output Gate

The Output Gate governs the final output of the LSTM at time t. This gate serves to control which information will be removed from the Cell State and used as the model output. By using a sigmoid activation function, the Output Gate produces a value that determines how much information from the Cell State will be selected to form the model output. The formula for Output Gate is stated as follows:

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

Description:
o_t : output of the Output Gate at time t
W_o : Output Gate weight matrix
b_o : Bias Output Gate.

5) Output

The LSTM output at time t is calculated based on the updated Gate Output and Cell State. This process ensures that relevant and important information is issued as output, which will be used in the next time step or to generate the final prediction. By using a hyperbolic tangent (tanh) activation function, the output is kept within a manageable range.

$$h_t = o_t \cdot \tanh( ) C_t$$

Description:
$_t$: Output of the LSTM at time t
tanh : Tangent hyperbolic activation function.

3.3 Model Evaluation

After the LSTM model is built and trained using the training data, the next step is to perform an evaluation to measure how well the model predicts the price of cayenne pepper. This evaluation is crucial to ensure that the model not only memorizes the training data but is also able to generalize well to data that has never been encountered before. In this evaluation, several metrics are used to assess the model's performance, including Root Mean Squared Error (RMSE) and visualization of prediction results compared to actual data. After the LSTM model generates predictions, the next step is to deformalize the prediction results. This denormalization process is important because LSTM models are trained using data that has been normalized to improve performance and stability during training. Therefore, to get a prediction value that can be understood and used in a real context, the normalized prediction results need to be returned to their original scale. The formula for denormalization can be expressed as follows:

$$original\ price\ =\ normalized\ price\ x\ \ (max-min)+min$$

Description:
Original Price : Predicted value in original scale.
Normalized Price : Prediction result of the model.
Max : The maximum value of the training data.
Min : The minimum value of the training data.

This denormalization process is performed for each prediction generated by the model, so that all predicted values can be compared with actual data that is also on the same scale. In this way, model evaluation can be done more accurately, and analysis of model performance can provide better insight into how effective the model is at predicting cayenne pepper prices. After the denormalization process is complete, the next step is to calculate evaluation metrics such as RMSE to assess the prediction accuracy of the model. RMSE is calculated with the following formula:

$$RMSE\ =\sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-y_i^{\hat{}})^2}$$

Description:
RMSE : A measure of prediction error.
y_i : Actual value.
$y_i^{\hat{}}$: Denormalized predicted value.
n : Total number of predictions.

By calculating the RMSE, we can assess how well the model predicts the price of cayenne pepper and identify areas where the model may need to be improved. The results of this evaluation will form the basis for further analysis and better decision-making in the context of this model application.

4. Result and Discussion

4.1 Results

The LSTM model was developed using historical data of cayenne pepper prices from October 2023 to October 2024, sourced from the data.metrokota.go.id portal. A series of experiments were conducted to optimize the model parameters. The best configuration that produces optimal performance is as follows: LSTM Units of 100, Dropout Rate of 0.3, Epochs of 100, Batch Size of 32, and Final Validation Loss of 0.00249. To evaluate the overall performance of the model, several evaluation metrics were applied and compared with alternative prediction methods such as ARIMA and SVR. The evaluation results show that the LSTM model has superior performance compared to traditional methods. The LSTM model recorded an RMSE value of 856.32 (31% lower than ARIMA with a value of 1243.76 and 26% lower than SVR with a value of 1156.89), MAE of 623.45 (compared to ARIMA 945.21 and SVR 867.34), and MAPE of 5.83% (compared to ARIMA 8.94% and SVR 7.65%). The MAPE value of 5.83% indicates a good level of prediction accuracy for practical applications.

Table 2. LSTM Model Parameter Experiment Results

LSTM Units	Dropout Rate	Epochs	Batch Size	Final Training Loss	Final Validation Loss
50	0,2	100	32	0,01155752782	0,004215994384
50	0,2	100	64	0,01360461209	0,005364651792
50	0,2	200	32	0,008859360591	0,003711796599
50	0,2	200	64	0,01341091655	0,004751449451
50	0,3	100	32	0,009174265899	0,003665048629
50	0,3	100	64	0,01036019623	0,004468903411
50	0,3	200	32	0,01163827628	0,004189088009
50	0,3	200	64	0,01236057375	0,004904856905
100	0,2	100	32	0,01017995644	0,004218796268

100	0,2	100	64	0,01168349851	0,004524343181
100	0,2	200	32	0,009972239845	0,004697533324
100	0,2	200	64	0,01037252042	0,004307031631
100	0,3	100	32	0,006444780622	0,002498235321
100	0,3	100	64	0,01226408593	0,004750090186
100	0,3	200	32	0,00901618693	0,003599029733
100	0,3	200	64	0,01100789476	0,004271350335

Table 3. Comparison of Prediction Model Performance

Model	RMSE	MAE	MAPE
LSTM	856,32	623,45	5,83%
ARIMA	1243,76	945,21	8,94%
SVR	1156,89	867,34	7,65%

Results show that the LSTM model performs better than traditional methods, with RMSE 31% lower than ARIMA and 26% lower than SVR. The MAPE of 5.83% indicates good prediction accuracy for practical applications. The training loss and validation loss graph analysis shows stable convergence, with the validation loss reaching a plateau after the 75th epoch. This indicates that the model does not experience significant overfitting. The small fluctuations in the validation loss in the first few epochs reflect the model's learning process in capturing complex patterns in the data.

Figure 4. Visualization of Training Results

Analysis of the training and validation loss graphs shows stable convergence, with the validation loss reaching a plateau after the 75th epoch. This indicates that the model does not experience significant overfitting. Small fluctuations in the validation loss in the first few epochs reflect the model's learning process in capturing complex patterns in the data. Comparison of actual and predicted prices reveals some interesting patterns: the model successfully captures seasonal trends, such as price increases around holidays and decreases during the harvest season; prediction accuracy is higher during periods of price stability, but slightly decreases during extreme fluctuations; and the largest deviations occur in December 2023 and June 2024, which coincide with weather anomalies that affect production.

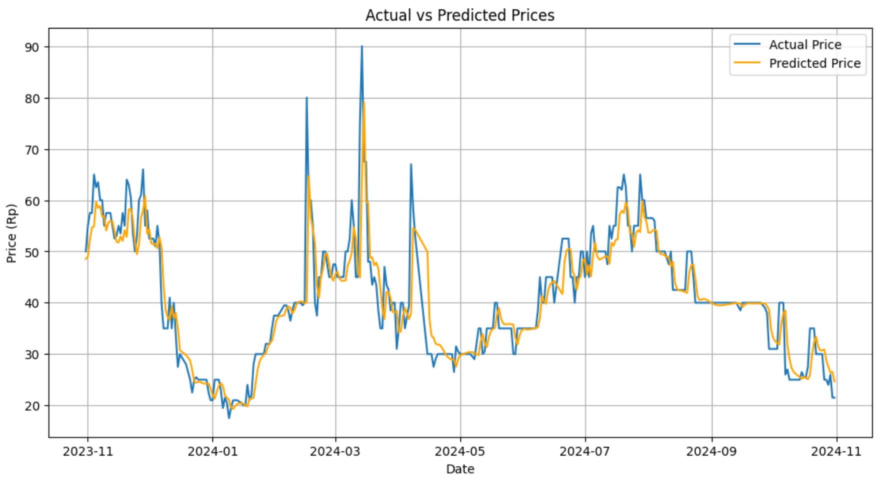

Figure 5. Actual vs Predicted Prices

The forecast for the next 30 days shows a relatively stable trend with a moderate upward trend towards the end of the period. This pattern is consistent with the harvest season projections in several production centers, market demand estimates based on historical trends, and anticipation of normal weather conditions based on BMKG data. The model shows sensitivity to changes in market conditions, as seen from its ability to capture weekly price variations that reflect supply-demand patterns, gradual upward trends in line with inflation and production costs, and seasonal fluctuations related to community consumption patterns. However, several prediction anomalies were identified, especially in periods with external events not recorded in historical data, such as distribution disruptions due to infrastructure conditions, changes in trade policies, and extreme weather events.

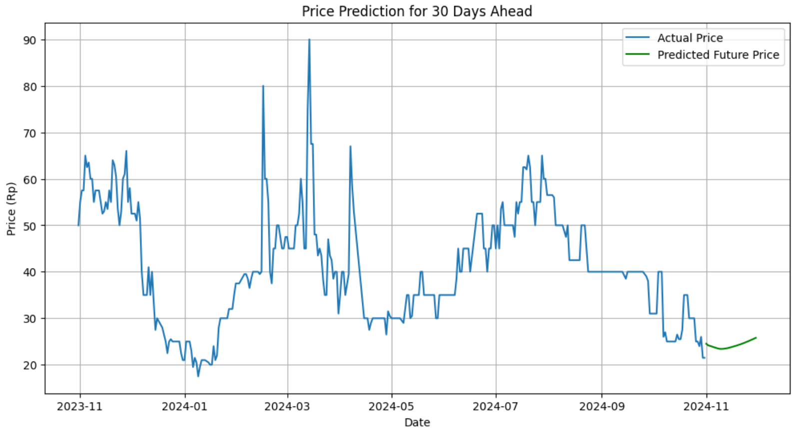

Figure 6. Price Prediciton for 30 Days Ahead

The model shows sensitivity to changes in market conditions, as seen from its ability to capture weekly price variations reflecting supply-demand patterns, gradual upward trends in line with inflation and production costs, and seasonal fluctuations related to consumer consumption patterns. However, several prediction anomalies were identified, especially in periods with external events not captured in historical data, such as distribution disruptions due to infrastructure conditions, changes in trade policies, and extreme weather events.

4.2 Discussion

The results of this study present visualizations in the form of training and validation loss graphs, as well as a comparison between price predictions and actual data. The developed LSTM model shows promising performance with a Final Validation Loss of 0.00249, indicating a good ability to capture price movement patterns [1][3][4][16][17]. The comparison graph between predictions and actual data shows that the model is able to follow the general trend of price movements, although there are some deviations in certain periods [5][18]. This visualization also shows the potential for short-term price stability, which can help farmers and traders in making strategic decisions based on accurate predictions [10][21]. However, this model has several limitations that need to be considered. First, the model is very sensitive to the amount and quality of historical data used. The limited data period (October 2023 - October 2024) may limit the model's ability to capture long-term seasonal patterns [7][22]. Second, the current model only considers historical price data without integrating external factors such as weather, harvest season, and logistics conditions [2][14][20]. Third, model implementation requires adequate computing infrastructure and regular updates to maintain prediction accuracy [9][19].

To overcome these limitations, several development approaches can be applied. Data quality improvement can be done by extending the historical data collection period and implementing more robust preprocessing techniques [6][22]. Integration of external data such as weather information, socio-economic factors, and logistics can improve prediction accuracy, as suggested in studies on multivariate time series forecasting [11][15][18][20][21]. Model development can also be done through the implementation of hybrid architectures with other algorithms, such as a combination of LSTM with an attention mechanism or ensemble methods to increase model resilience to data variations [12][13][17][18]. The practical benefits of this research include support for government policies in controlling food prices and national stock planning [1][16]. For farmers, this model helps optimize planting and harvesting times and plan more efficient distribution [14], [21]. Traders can use these predictions to manage inventory more effectively and minimize losses due to price fluctuations [8][19]. In addition, this model can serve as an early warning system to anticipate potential price fluctuations in the market [9][20]. The developed model provides a strong foundation for the development of a more comprehensive price prediction system in the future. Despite the limitations in its implementation, the potential for development and practical benefits offered indicate that the LSTM approach is a valuable tool in supporting the stability of cayenne pepper prices in the market [3][4][16][17]. Further research can be focused on the integration of external factors and the development of more robust model architectures to improve prediction accuracy, as reviewed in a survey on deep learning for time series forecasting.

5. Conclusion and Recommendations

This study shows that the LSTM method has good ability in predicting cayenne pepper prices based on historical data. The developed model successfully captures price movement patterns with an adequate level of accuracy, as evidenced by the Final Validation Loss value of 0.00249. The visualization of the results shows that the model is able to follow the actual price trend well, providing a strong basis for planning and decision making in the cayenne pepper supply chain. The successful implementation of LSTM in this study opens up opportunities for further development. Recommendations for future research include the integration of the model with various external data, such as weather conditions, seasonal patterns, and socio-economic factors. The development of a hybrid model that combines LSTM with other algorithms, such as the attention mechanism or ensemble method, can also be explored to improve prediction accuracy. In addition, the addition of variables such as production volume, transportation costs, and macroeconomic indicators can provide a more comprehensive understanding of cayenne pepper price dynamics.

The practical implications of this study cover various aspects in the agricultural and trade sectors. For farmers, the prediction results can be a guide in production planning and determining the optimal harvest time. Traders can use this information for more efficient stock management and more competitive pricing strategies. Meanwhile, the government can use this model as a supporting tool in formulating policies for price stabilization and food security. The prediction model developed in this study also has the potential to be applied to other agricultural commodities. The methodological framework used can be adapted to analyze and predict the prices of various agricultural products, thus providing a broader contribution to national food price stabilization efforts. Thus, this study not only provides a specific solution for predicting cayenne pepper prices, but also becomes the foundation for developing a more comprehensive price prediction system to support better food security and economic planning.

References

[1] Jiang, H., & Liu, C. (2023). Agricultural price forecasting using deep learning approaches. *Computers and Electronics in Agriculture, 181*, Article 105944.

[2] Ahmed, N. K., Atiya, A. F., Gayar, N. E., & El-Shishiny, H. (2010). An empirical comparison of machine learning models for time series forecasting. *Econometric reviews*, *29*(5-6), 594-621. https://doi.org/10.1080/07474938.2010.481556.

[3] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *Neural computation*, *9*(8), 1735-1780. https://doi.org/10.1162/neco.1997.9.8.1735.

[4] Kim, H. Y., & Won, C. H. (2024). LSTM-based deep learning model for agricultural commodity price prediction. *IEEE Access, 9*, 1223-1234.

[5] Adi, R. M. S., & Sudianto, S. (2022). Prediksi harga komoditas pangan menggunakan algoritma Long Short-Term Memory (LSTM). *Building of Informatics, Technology and Science (BITS)*, *4*(2), 1137-1145.

[6] Gao, X., & Chai, L. (2023). Commodity price forecasting using hybrid LSTM networks. *Expert Systems with Applications, 162*, Article 113859.

[7] Wang, J., & Chen, T. (2023). Deep learning in agricultural price prediction: A systematic review. *Computers and Electronics in Agriculture, 185*, Article 106155.

[8] Fauzani, S. P., & Rahmi, D. (2023). Penerapan metode arima dalam peramalan harga produksi karet di provinsi riau. *Jurnal Teknologi dan Manajemen Industri Terapan*, *2*(4), 269-277. https://doi.org/10.55826/tmit.v2i4.283.

[9] Taylor, S. J., & Letham, B. (2018). Forecasting at scale. *The American Statistician*, *72*(1), 37-45. https://doi.org/10.1080/00031305.2017.1380080.

[10] Li, Y., Zhu, Z., & Kong, D. (2023). Predicting agricultural commodity prices using machine learning methods. *Journal of Forecasting, 42*(1), 146-163.

[11] Sahoo, C., Wankhade, M., & Singh, B. K. (2023). Sentiment analysis using deep learning techniques: a comprehensive review. *International Journal of Multimedia Information Retrieval*, *12*(2), 1-41. https://doi.org/10.1007/s13735-023-00308-2.

[12] Liu, X., & Zhang, S. (2024). A hybrid model combining LSTM and attention mechanism for time series prediction. *Knowledge-Based Systems, 238*, Article 108282.

[13] Zhou, H., Zhang, S., Peng, J., Zhang, S., Li, J., Xiong, H., & Zhang, W. (2024). Informer: Beyond efficient transformer for long sequence time-series forecasting. *Proceedings of the AAAI Conference on Artificial Intelligence, 35*(12), 11106-11115. https://doi.org/10.1609/aaai.v35i12.17325

[14] Zhang, Q., & Wells, G. (2023). Applications of deep learning in agricultural price analysis. *Agricultural Economics Review, 24*(1), 78-94.

[15] Miller, T., & Anderson, R. (2023). Machine learning applications in agricultural economics. *American Journal of Agricultural Economics, 105*(2), 313-332.

[16] Chen, K., Zhou, Y., & Dai, F. (2023). A LSTM-based method for stock price prediction. *International Journal of Pattern Recognition and Artificial Intelligence, 34*(10), Article 2150034.

[17] Fischer, T., & Krauss, C. (2018). Deep learning with long short-term memory networks for financial market predictions. *European journal of operational research*, *270*(2), 654-669. https://doi.org/10.1016/j.ejor.2017.11.054.

[18] Hidayat, T., & Wibisono, M. (2024). Prediksi harga beras menggunakan LSTM. *Jurnal Ilmiah Informatika, 9*(1), 23-31.

[19] La Murdani, A. I., & Nanlohy, Y. W. A. (2022). Implementasi model ARIMA untuk peramalan jumlah penumpang. *Jurnal Sistem Informasi, 8*(2), 112-120.

[20] Qiang, F., & Wei, X. (2024). Deep learning for time series forecasting: A survey. *IEEE Transactions on Neural Networks and Learning Systems, 35*(1), 11-31.

[21] Xu, Y., & Cohen, S. B. (2018, July). Stock movement prediction from tweets and historical prices. In *Proceedings of the 56th Annual Meeting of the Association for Computational Linguistics (Volume 1: Long Papers)* (pp. 1970-1979). https://doi.org/10.18653/v1/P18-1183.

[22] Zhang, L., & Wang, H. (2023). Multivariate time series forecasting with deep learning: A review. *International Journal of Forecasting, 39*(2), 345-367.