ISSN : 1978-8282, Online ISSN: 2655-4275

Comparative Analysis of Time Series Methods LSTM and ARIMA for Predicting Inventory Availability (Case Study: PT XYZ)

Edi Kartawijaya[*1], Munawar[2], Gerry Firmansyah[3], Budi Tjahjono[4]
[1,2,3,4] Master of Computer Science Study Program, Faculty of Computer Science, Esa Unggul University Jakarta, Indonesia
E-mail: [*1]**windikster@gmail.com**, [2]moenawar@gmail.com, [3]gerry@esaunggul.ac.id, [4]budi.tjahjono@esaunggul.ac.id

Abstract

Product availability plays a crucial role in supply chain management, directly impacting all aspects of business operations, from production to distribution. This study analyzes the optimization of product availability at PT. XYZ, a frozen and chilled food trading company in Indonesia, focusing on four main commodities: beef, buffalo meat, chicken, and potatoes. Utilizing historical transaction data from 2020 to July 27, 2024, this research compares the performance of two forecasting models: ARIMA (AutoRegressive Integrated Moving Average) and Long Short-Term Memory (LSTM), in predicting product availability The traditional ARIMA model has proven effective in time series data analysis but has limitations in capturing complex patterns and non-linear fluctuations. LSTM, as a machine learning technique, demonstrates superiority in capturing long-term temporal relationships. This study finds that the LSTM model consistently outperforms ARIMA for beef, buffalo meat, and chicken categories, although there is a slight increase in error for the potatoes category. Model performance evaluation is conducted using metrics such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE). The results indicate that the LSTM model exhibits lower errors compared to ARIMA, proving its effectiveness in predicting dynamic demand patterns. With a better understanding of product availability, the company is expected to reduce operational costs, avoid losses, and enhance customer satisfaction through more efficient supply chain management. This research provides significant insights for PT. XYZ and similar industries in implementing more accurate forecasting methodologies.

Keywords *— ARIMA, LSTM and Predicting*

1. INTRODUCTION

Inventory availability plays a critical role in supply chain management as it affects all aspects of business operations from production to distribution. Optimizing inventory through advanced forecasting models can create a competitive advantage[1]. For companies managing products like beef, buffalo meat, chicken, and potatoes, operational costs are significantly impacted by storage requirements. These products necessitate special facilities, such as energy-intensive refrigeration, to maintain quality [2].

Traditional forecasting models like ARIMA have been effective for time series analysis but fall short in capturing complex, nonlinear patterns. Machine learning approaches, particularly Long Short-Term Memory (LSTM), have emerged as promising alternatives due to their ability to capture long-term temporal dependencies, providing more accurate predictions in dynamic environments [3]. Studies indicate that machine learning algorithms, including LSTM, significantly reduce prediction errors compared to classical statistical models [4], [5].

Evaluating model performance through metrics such as Mean Absolute Error (MAE), Mean Absolute Percentage Error (MAPE), and Root Mean Squared Error (RMSE) allows for a comprehensive assessment. These metrics help to quantify prediction accuracy and reveal significant prediction[5].

Inaccurate demand forecasting can lead to overstocking or stockouts, resulting in significant business losses [6]. Effective forecasting models are essential for companies like PT XYZ, a major player in Indonesia's cold and frozen food market, which faces challenges due to variable demand patterns. This study aims to compare ARIMA and LSTM to enhance inventory management. Using historical transaction data from PT XYZ from January 2020 to June 2024, it focuses on predicting availability for high-demand items like beef, buffalo meat, chicken, and potatoes.

2. RESEARCH METHOD

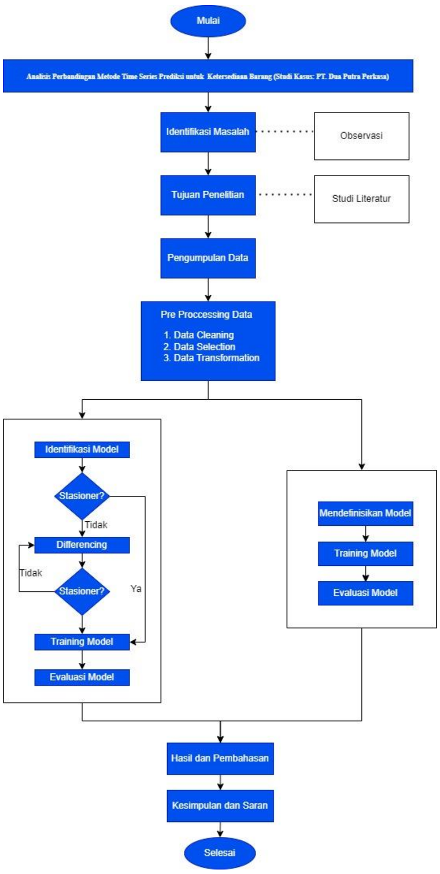

Figure 1. Research framework

In this study, several stages were conducted as shown in the diagram above. The research begins with problem identification and then determining the research objective, which is to compare forecasting methods for predicting inventory availability in terms of weight volume (kg).

Following this, data is collected from PT XYZ's internal database, with details of the data explained in section 3.2. To develop a robust model, the collected data must first undergo preprocessing to be suitable for forecasting. Preprocessing includes data cleaning, data selection, and data transformation (converting data format). Once the data has been preprocessed, modeling is conducted using two different methods: ARIMA and LSTM.

In the ARIMA modeling phase (branch on the left), several steps are taken prior to modeling the data. These steps include checking whether the data is stationary to meet the assumption that the data must be stationary. If the data is stationary, it can be directly modeled using the ARIMA model. If not, differencing techniques are applied until the data becomes stationary before modeling. Model evaluation is then performed using evaluation metrics such as MAE, RMSE, and MAPE.

In contrast, the LSTM modeling phase (branch on the right) does not require checking data stationarity. Instead, the process involves model identification with certain hyperparameters, followed by model training and ending with model evaluation using the same metrics as the ARIMA model. After obtaining results from both ARIMA and LSTM models, a comparison is made based on measurable evaluation metrics and forecasting results.

The study concludes with a determination of which model is best for predicting inventory availability at PT XYZ and provides recommendations for future development.

3. RESEARCH RESULTS AND DISCUSSION

A. Dataset

The data used in this study is sourced from transaction data of PT XYZ from January 2020 to June 2024. Below are the available columns along with their descriptions:

Nama Kolom	Non Null Count	Tipe Data
Invoice No.	267251	object
Invoice Date	267251	datetime64[ns]
Item No.	267251	object
Item Description	267251	object
Quantity	267251	float64
Item Unit 1	214760	object
Customer Name	267251	object
Salesman Name	267251	object
project	267251	object

Nama Kolom	Non Null Count	Tipe Data
Month	267251	object
Group Customer	267004	object
Tahun	267251	int64
Group	267251	object
Jenis	267251	object
Type	267251	object
Origin	267251	object
Merk	267236	object
Kriteria	267251	object
New Code Marketing	267251	object
Marketing	267208	object
Group2	267251	object

B. ARIMA Analysis Results

Based on the discussion in Section 3.4.1, before the data was modeled, stationarity testing, differencing, and ACF and PACF analysis were conducted to determine the parameter values of p, d, and q for each product category. The values for p, d, and q for each category based on the completed steps are displayed below.

Kategori	p	d	q
Kentang	2	1	2
Daging Sapi	1	0	1
Daging Kerbau	1	0	1
Ayam	1	0	1

The researchers applied the above parameters to each model per category and obtained evaluation metrics results, including MAE, RMSE, and MAPE, as shown in the table below.

Kategori	RMSE	MAE	MAPE
Kentang	92504.54	86005.89	1.34%
Daging Sapi	432253.76	360222.34	0.26%
Daging Kerbau	283796.16	215029.96	0.31%
Ayam	23507.78	21434.87	0.88%

In theory, RMSE and MAE values are considered good when they are close to zero. The table above shows that no category has RMSE and MAE values close to zero; in fact, the values are quite high. This is expected, as RMSE and MAE represent the difference or error between the predicted and actual values. These error values should be compared with the range of data values to determine if the error is significant or not. Based on the MAPE value, which provides this information, the percentage error for all categories is very small, even below 2%.

Based on the evaluation metric values, it can be concluded that the ARIMA model for all four categories can be used for forecasting modeling.

C. LSTM Analysis Results

The LSTM modeling was applied to all four data categories using the same model parameters and resulted in models with the following evaluation metric values.

Kategori	RMSE	MAE	MAPE
Kentang	100160.41	92569.31	1.37%
Daging Sapi	414213.49	330155.33	0.21%
Daging Kerbau	249154.76	178167.50	0.17%
Ayam	17974.13	13483.31	0.37%

Based on the evaluation metrics above, the LSTM model produced better evaluation metrics for the categories of beef, buffalo meat, and chicken. In the potato category, the evaluation metrics improved, but not significantly.

For the LSTM model, it is necessary to assess the model's performance during training, as RMSE, MAE, and MAPE metrics alone are not sufficient to fully describe how the model will perform in predicting future data. Therefore, analyzing the model's performance through the loss value graph against epochs during the model training process is essential.

D. Comparison of Forecasting Results

In sub-sections 4.2 and 4.3, the discussion focuses on the model's performance in predicting data, specifically training and test data. We can observe that the error values remain within a small range, with MAPE below 2%. However, to determine whether the model can truly predict data effectively, we need to delve deeper rather than relying solely on evaluation metrics that only provide average values. Evaluation metrics do not reveal trends or other time series components. The following section will discuss how the model can perform forecasting for at least the training and test data.

3.1. Potato

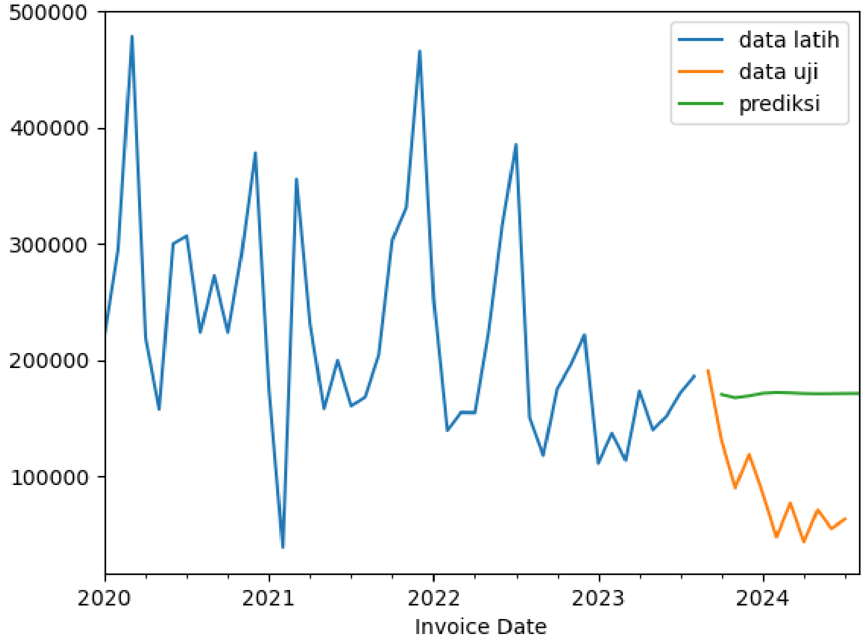

Figure 2. The predicted graph of the ARIMA model for the Potato category.

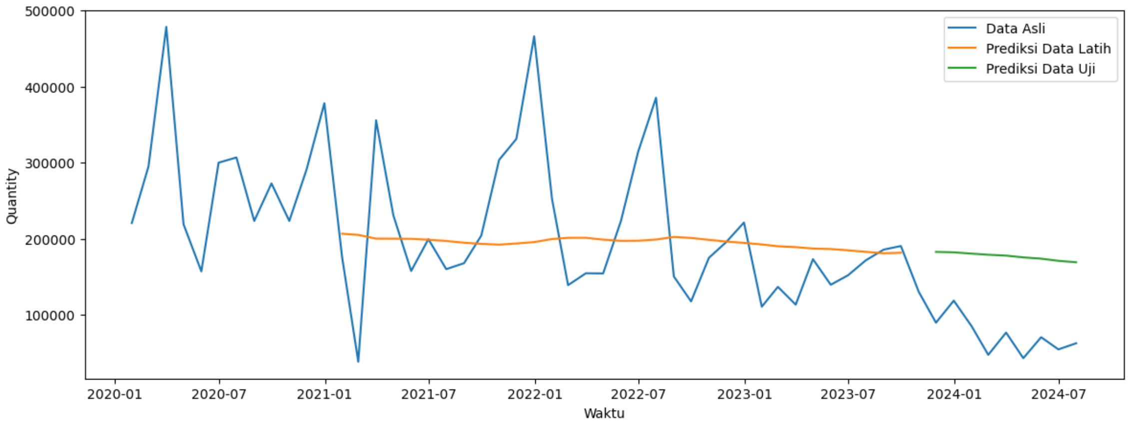

Figure 3. The predicted graph of the LSTM model for the Potato category.

In Graphs 5 and 6, the forecasting results of the ARIMA and LSTM models for the potato category are shown. The colors in Graph 5 represent the original data, which includes training data (blue), original test data (orange), and predicted data for the test set (green). Graph 6 contains the original data (blue), training data predictions based on the last 12 data points/year (orange), and predictions for the test data (green).

Both graphs indicate that neither model has successfully predicted the data accurately, despite capturing trend patterns. For the LSTM model, this was anticipated earlier in the discussion of sub-section 4.2, where it was noted that the model is less stable compared to other categories.

3.2. Daging Sapi

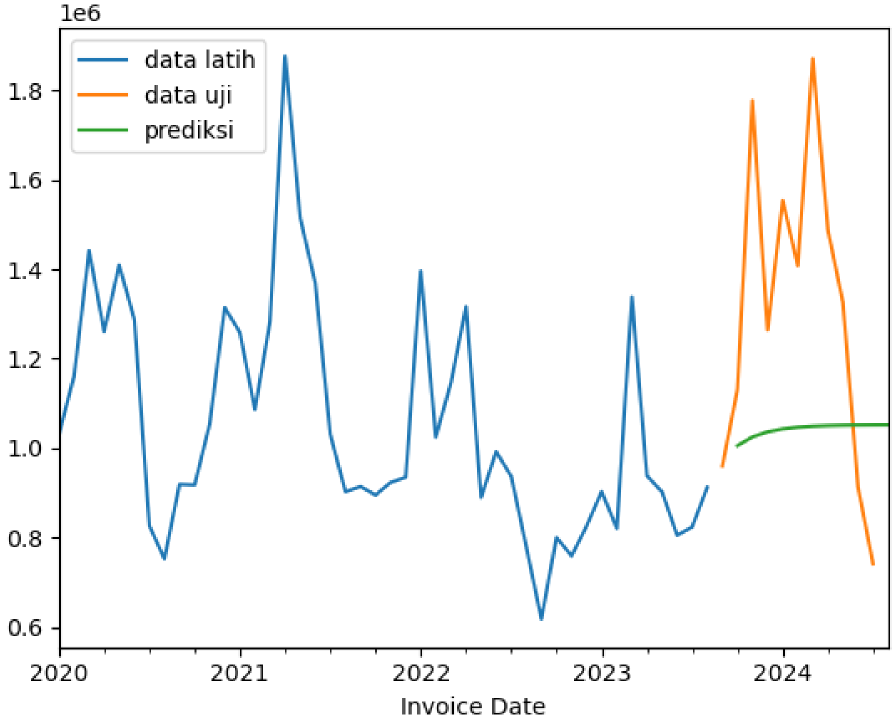

Figure 4. The predicted graph of the ARIMA model for the Beef category.

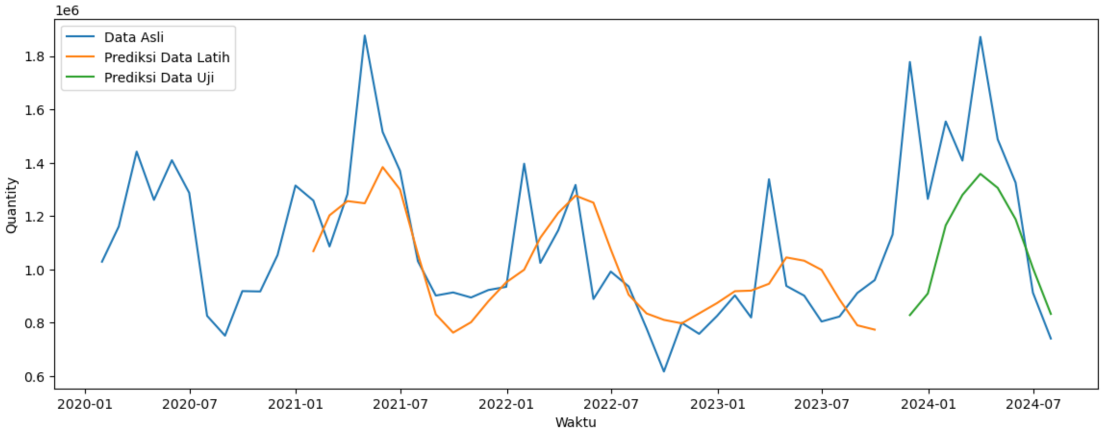

Figure 5. The predicted graph of the LSTM model for the Beef category.

For the beef category model, Graph 7 shows that the ARIMA model is not yet able to predict the data well, even though the evaluation metrics, on average, are very good. In contrast, Graph 8 indicates that while the exact prediction figures are quite far from the actual values, the predicted results can capture the trends and seasonal patterns well.

3.3. Buffalo Meat

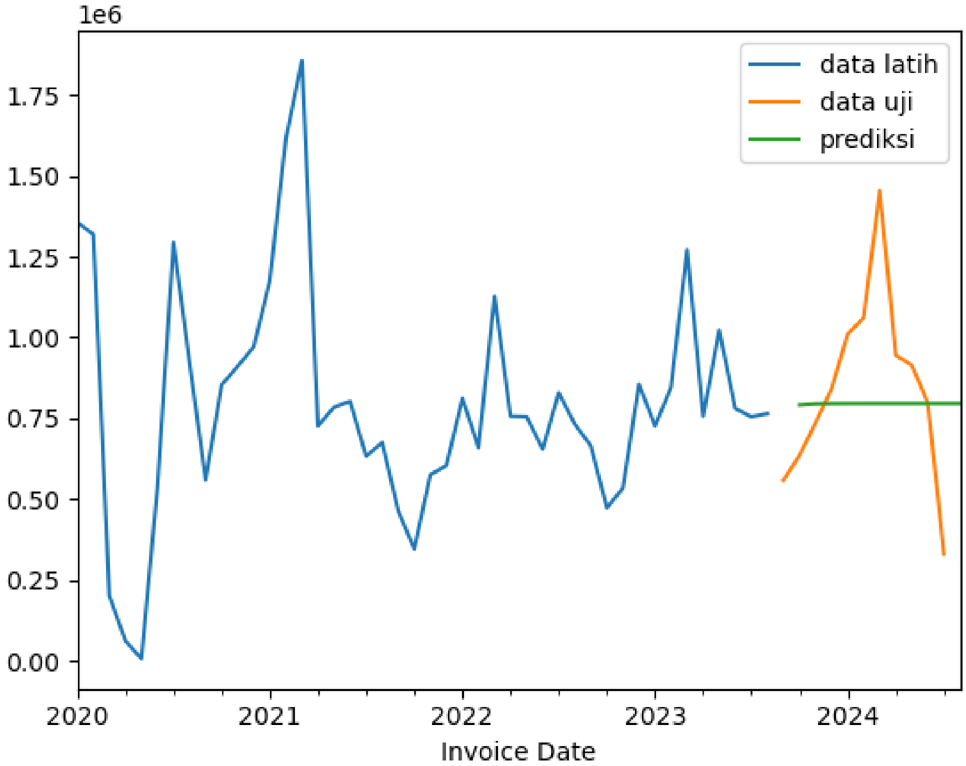

Figure 6. The predicted graph of the ARIMA model for the Buffalo Meat category.

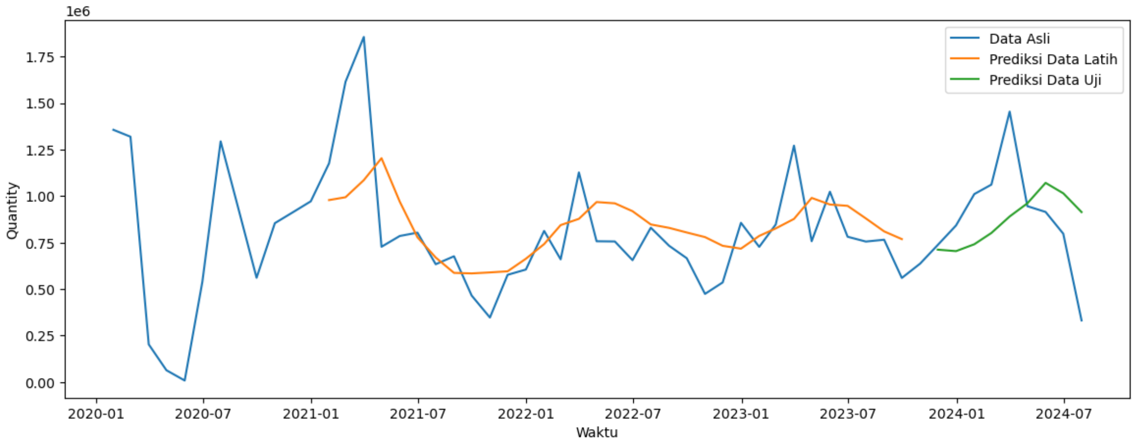

Figure 7. The predicted graph of the LSTM model for the Buffalo Meat category.

The model for the buffalo meat category also shows similar results to beef, where the data patterns can be followed well by the LSTM model for both training and test data. However, the ARIMA model does not show good results.

3.4. Chicken

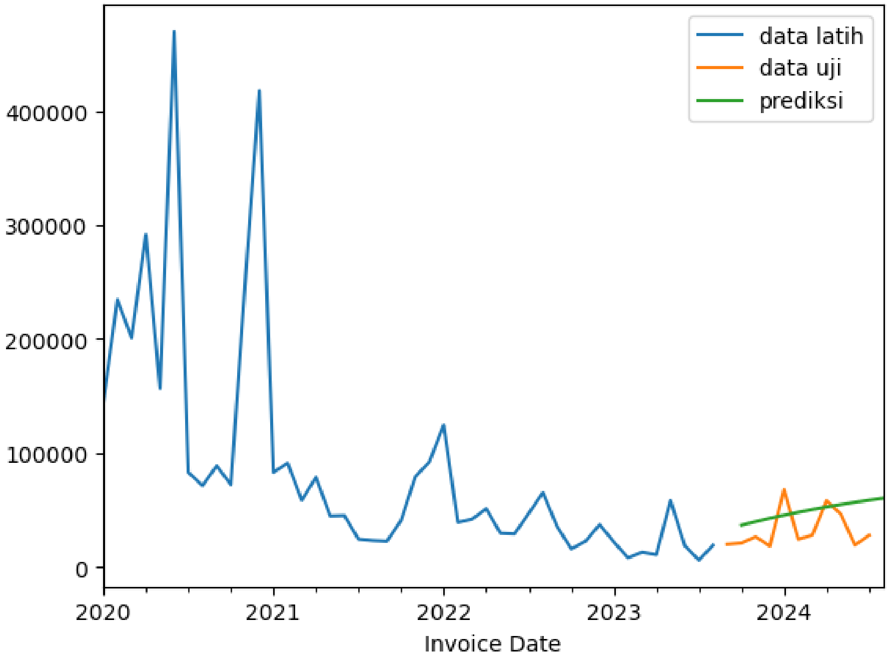

Figure 8. The predicted graph of the ARIMA model for the Chicken category.

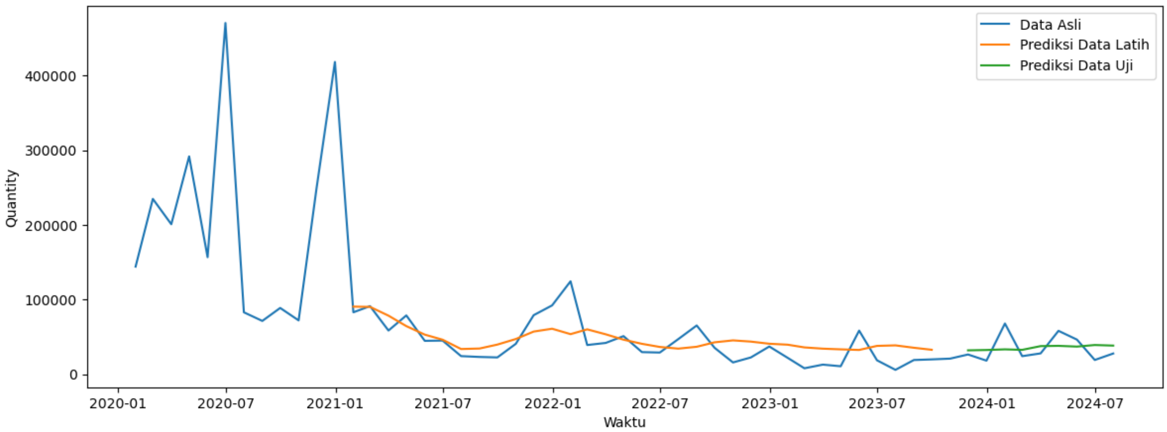

Figure 9. The predicted graph of the LSTM model for the Chicken category.

In the model for the chicken category, trend patterns can be followed/predicted fairly well by both the ARIMA and LSTM models. However, the details of the data's increases and decreases are not quite accurate due to the initial data, which has a different pattern compared to the rest.

Table 1. Perbandingan hasil prediksi model LTSM dengan model ARIMA

Kategori	ARIMA			LSTM		
	RMSE	MAE	MAPE	RMSE	MAE	MAPE
Kentang	92504.54	86005.89	1.34%	100160.41	92569.31	1.37%
Daging Sapi	432253.76	360222.34	0.26%	414213.49	330155.33	0.21%
Daging Kerbau	283796.16	215029.96	0.31%	249154.76	178167.50	0.17%
Ayam	23507.78	21434.87	0.88%	17974.13	13483.31	0.37%

The table above illustrates the comparison of prediction results between the LSTM model and the ARIMA model. This research shows that the LSTM model results are better than ARIMA for the categories of Beef, Buffalo Meat, and Chicken. For the Potato category, the error rate increased compared to ARIMA, but not significantly. Based on the evaluation metrics RMSE, MAE, and MAPE, both models demonstrate good performance in predicting inventory availability. The accuracy rate for the ARIMA method is 99.30% and for LSTM is 99.47%.

Table 2. Hasil Mean Absolute Percentage Error (MAPE)

Metode	MAPE (Ayam)	MAPE (Daging Sapi)	MAPE (Daging Kerbau)	MAPE(Kentang)	Grand Total (Ayam,Daging Sapi,Kerbau,Kentang)
ARIMA	0,88%	0,26%	0,31%	1,34%	
LSTM	0,37%	0,21%	0,17%	1,37%	
Biaya (ARIMA) Periode Januari 2020-Juni 2024					
Storage	Rp270.585.558	Rp1.152.330.709	Rp1.003.225.347	Rp1.055.813.763	Rp3.481.955.377
Handling	Rp2.594.656	Rp766.603	Rp914.027	Rp3.950.954	Rp8.226.239
Biaya (LSTM) Periode Januari 2020-Juni 2024					
Storage	Rp113.768.928	Rp64.571.554	Rp52.272.210	Rp421.252.516	Rp651.865.207
Handling	Rp1.090.935	Rp619.179	Rp501.240	Rp4.039.408	Rp6.250.762
Biaya (ARIMA) Rata-rata per tahun					
Storage	Rp60.130.124	Rp256.073.491	Rp222.938.966	Rp234.625.281	Rp773.767.861
Handling	Rp576.590	Rp170.356	Rp203.117	Rp877.990	Rp1.828.053
Biaya (LSTM) Rata rata per tahun					
Storage	Rp25.281.984	Rp14.349.234	Rp11.616.047	Rp93.611.670	Rp144.858.935
Handling	Rp242.430	Rp137.595	Rp111.387	Rp897.646	Rp1.389.058

The table above is derived from the Mean Absolute Percentage Error (MAPE), storage/handling costs, and quantity of each category to calculate the value of storage and handling costs. After calculation, the company can maximize potential figures based on the ARIMA model with an annual average for storage of Rp773,767,861 and handling of Rp1,828,053, whereas for the LSTM model, storage is Rp144,858,935 and handling is Rp1,389,058.

4. CONCLUSION

This research shows that the LSTM model results for the categories of Beef, Buffalo Meat, and Chicken are better than those of ARIMA. For the Potato category, the error increased compared to ARIMA but not significantly. Based on the evaluation metrics RMSE, MAE, and MAPE, both models demonstrate good performance in predicting inventory availability. However, LSTM proved to be superior in this analysis, showing lower error values across all three metrics. The accuracy rate for the ARIMA method is 99.30% and for LSTM is 99.47%.

The forecasting results indicate that LSTM excels due to its ability to predict the upward and downward patterns of data more closely resembling the actual data. Company XYZ can maximize potential figures based on the ARIMA model with an annual average for storage of Rp773,767,861 and handling of Rp1,828,053, whereas for the LSTM model, storage costs are Rp144,858,935 and handling costs are Rp1,389,058.

5. REFERENCES

[1] L. N. K. Duong and J. Chong, "*Supply chain collaboration in the presence of disruptions: a literature review*," Int J Prod Res, vol. 58, no. 11, pp. 3488–3507, 2020.

[2] W. Al-Zaidi, A. Al-Karawi, and A. Al-Zuhairi, "*Role of Demand Forecasting and Lead Time on Waste in Supply Chain: A Case Study in Diyala Health Sector-Iraq*," Journal of Advanced Management Science, pp. 155–160, Aug. 2018, doi: 10.18178/joams.6.3.155-160.

[3] O. B. Sezer, M. U. Gudelek, and A. M. Ozbayoglu, "*Financial time series forecasting with deep learning: A systematic literature review: 2005–2019*," Appl Soft Comput, vol. 90, p. 106181, 2020.

[4] S. N. Mohanty, P. Nanjundan, and T. Kar, "*Artificial Intelligence in Forecasting*".

[5] A. Botchkarev, "*Performance metrics (error measures) in machine learning regression, forecasting and prognostics: Properties and typology*," arXiv preprint arXiv:1809.03006, 2018.

[6] A. Al Mamun, M. Sohel, N. Mohammad, M. S. H. Sunny, D. R. Dipta, and E. Hossain, "*A comprehensive review of the load forecasting techniques using single and hybrid predictive models*," IEEE access, vol. 8, pp. 134911–134939, 2020.