Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Performance Comparison of ARIMA, LSTM, and Prophet Methods in Sales Forecasting

I Gede Totok Suryawan[1)*], I Kadek Nurcahyo Putra[2)], Putu Mita Meliana[3)], I Gede Iwan Sudipa[4)]
[1,2, 3,4)]Department of Informatics, Institute of Business and Technology Indonesia, Denpasar, Bali, Indonesia
[1)*]totok.suryawan@gmail.com , [2)]nurcahyo@instiki.ac.id , [3)]mitha.mitha789@gmail.com ,
[4)]iwansudipa@instiki.ac.id

Sent : Aug 21, 2024 | **Accepted** : Sep 28, 2024 | **Published** : Oct 6, 2024

Abstract: The development of the business world that is growing rapidly today resulted in tighter competitiveness between fellow business actors. One of the businesses that has sprung up in the market today is the bakery business. Currently, bread is one of the food needs in Indonesia that is great demand by children to the elderly, which is often used as breakfast or snack. One of the companies that produces white bread is the Bandung White Bread Factory. The number of sales at this factory continues to increase every month based on total sales data recorded since 2021. With the increasing number of sales at this factory, the factory often experiences stock shortages and cannot meet customer demand. Therefore, in this study, a model has been developed to forecast the sales of white bread using the ARIMA, LSTM, and Prophet methods. The results of the study showed that the ARIMA method (1,0,2) had the best performance compared to the LSTM and Prophet methods, because the ARIMA method (1,0,2) produced the smallest error accuracy value, namely with a MAPE value of 4.548%, an MSE value of 2248.0822, and an RMSE value of 47.4139.

Keyword : *Forecasting; Sales; ARIMA; LSTM; Prophet*

INTRODUCTION

The development of the business world is currently growing rapidly, especially in product sales, which causes competition between business actors to be tighter. One of the rapidly growing business sectors is the food industry, including bread (Nurarofah et al., 2023). Nowadays, bread is one of the basic needs that is in great demand by various groups of people in Indonesia, ranging from children to adults. White bread, known as one of the most common types of bread, is often consumed as breakfast or a snack (Makkulawu, 2023).

One of the companies that plays an important role in meeting the needs of white bread in Indonesia is Bandung White Bread Factory located in Denpasar, Bali. Established in 2020, the factory has recorded a consistent increase in white bread sales from March 2021 to May 2023. Every day, the factory is able to produce between 700 to 1100 packs of white bread. Nonetheless, with increasing demand, the factory often experiences stock shortages resulting in an inability to optimally fulfill consumer demand. On the other hand, there are also occasional overstocks, where products are not sold on time, which results in a decrease in the company's revenue.

These problems indicate an urgent need for factories to optimize stock management by predicting future bread sales. The purpose of this research is to develop and compare white bread sales forecasting models at the Bandung White Bread Factory using three main methods, namely ARIMA (Autoregressive Integrated Moving Average), LSTM (Long Short-Term Memory) and Prophet. These three methods were chosen for their ability to process complex time series data and produce accurate predictions (Ferdinandus et al., 2023; Sudipa et al., 2023; Yang, 2019).

In this study, daily sales data from March 2021 to May 2023 is used as the basis for forecasting. Each method will be tested using three forecasting accuracy measurements, namely Mean Absolute Percentage Error (MAPE), Mean Squared Error (MSE), and Root Mean Square Error (RMSE). The purpose of this test is to find out which method gives the lowest error rate (Suryadana & Sarasvananda, 2024; Suryawan et al., 2023), so that it can be used as the best model in sales forecasting at the bakery.

The implications of the research are particularly in optimizing inventory management and improving production efficiency. With an accurate forecasting model, the factory can better predict sales fluctuations, thereby reducing the risk of shortages and excess stock.

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

LITERATURE REVIEW

Several studies related to this theme have been conducted previously, including by Muhammad Ali Ridla, Nur Azise, and Miftahur Rahman (2023) in their research entitled "Comparison of Time Series Forecasting Models in Predicting the Number of Tourist Arrivals and Airport Passengers". The object of this study is the number of visitors to Taiwan and the number of passenger arrivals at Changi Airport Singapore. The purpose of this study is to predict the number of tourists in the future and to determine the performance of the SARIMA, LSTM, and Prophet methods through RMSE and MSE test measurements. From the results of forecasting passenger arrivals at Changi Airport Singapore using the three methods, it was found that the SARIMA method had the best performance compared to the other 2 methods with an RMSE value of 153,142 and an MSE of 23,452,430,508 on the data on the number of passenger arrivals at Changi Airport, and an RMSE value of 565 and an MSE of 319,425 on the data on the number of visitors to Taiwan (Ridla et al., 2023).

The second study was conducted by Jamaludin and Toto Haryanto (2023) in their study entitled "Utilization of the Long Short Term Memory (LSTM) Model to Predict Gold Prices as an Investment Instrument in Preparing for the Threat of a Global Recession in 2023". The purpose of this forecast is to predict gold prices in the future based on past data. This forecast was carried out using daily sales data (Monday-Friday) from November 16, 2017 to November 15, 2022. After conducting the study, the results were obtained with an RMSE accuracy value of 20, that the price of gold in 2023 will increase even though in reality there will be price fluctuations, so this method can be said to be robust and can be used to support decisions for investors in investing in gold (Haryanto, 2023).

The third study is a study conducted by Agnes Cherrly and Ramos Somya (2023) entitled "Prediction of Ticket Sales for Playground Tourism Using the ARIMA Method". The purpose of this study is to predict the number of visitors to the Saloka Theme Park Playground Tourism and minimize or anticipate queues at the get checking ticket or ticket purchase counter. This forecast was carried out using the recapitulation data for Saloka Theme Park ticket sales from January 2020 to January 2023. From the results of the study, it was found that the ARIMA model that is suitable for predicting the number of ticket sales for the Saloka Theme Park Playground Tourism is the ARIMA model (1,0,0) with an RMSE value of 21296.39 and an MSE value of 453536460.26 (Cherrly et al., 2023).

The fourth study is a study conducted by Iga Dwi Wahyuni, Trisna Yuniarti, and Amrin Rapi (2022) in their study entitled "Application of ARIMA Model in Predicting Sales of Sosro Bottled Tea Beverage Products Size 350 mL". The purpose of this forecasting is to obtain the best ARIMA forecasting model and compare it with the Moving Average method currently used by the company in predicting sales volume of TBE 350 mL K12 Aseptic products. This forecasting was carried out using sales data for the period January 2016-January 2022. From this study, the results showed that the best forecasting method was the ARIMA method (2,1,0) with an ME value of 624, MAD of 624, MSE of 389376, and MAPE of 2.690% (Wahyuni et al., 2022).

Another study was conducted by Yudha Alif Auliya, Yanuar Nurdiansyah, and Ari Puji Astuti (2023) entitled "Forecasting the Number of Visitors to the Gumul Paradise Island Tourist Attraction, Kediri Regency Using the Prophet Method". The purpose of this study was to predict and measure the level of accuracy of forecasting the number of visitors to the Gumul Paradise Island tourist attraction using the Prophet algorithm. The data used in this forecast is visit data for the last 10 years, namely from 2012 to 2022. From the study, it was found that forecasting using the Prophet method has a very good ability to predict the number of visitors to the Gumul Paradise Island tourist attraction, with a MAPE value of 9.758% in the forecast data compared to the original data 2022 (Auliya et al., 2023).

Based on several studies that have been conducted, in this study the author uses total sales data for white bread at the Bandung White Bread Factory from March 2021 to May 2023, using the ARIMA, LSTM, and Prophet methods which will later be tested using the MAPE, MSE, and RMSE validation tests, to find out which method can provide results with the smallest error rate in forecasting the number of sales at the Bandung White Bread Factory.

Autoregressive Integrated Moving Average (ARIMA)

ARIMA or what is often called the Box-Jenkins time series method is a model that combines the Autoregressive (AR) model, the Moving Average (MA) model, and the Autoregressive Moving Average (ARMA) model which is then developed into ARIMA by adding a parameter d or differencing (Alim, 2023). ARIMA is a model that completely ignores independent variables in making a forecast (Zuhri & Nisa, 2022). The accuracy of ARIMA is very good when used for short-term forecasting, whereas for long-term forecasting the accuracy of the forecast is not good and usually tends to be flat or horizontal (Lin, 2024; Sudipa et al., 2023). The ARIMA notation is p, d, q which indicates the processing order, p is the order for the autoregressive (AR) process, d represents the order which states the number of differencing processes, and q represents the order of moving average (MA) processing (Hayadi et al., 2021; Nurhasanah & Dini, 2023). The mathematical form of the ARIMA model is as follows:

$$Zt = \phi 1 Zt{-}s + \phi 2 Zt{-}2s + \cdots + \phi p Zt{-}ps + \alpha t - \theta 1 \alpha t{-}s - \theta 2 \alpha t{-}2s - \cdots - \theta q \alpha t{-}q\mathrm{s} \quad (1)$$

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

description:
Zt : stationary time series values,
$\alpha t-s$, $\alpha t-2s$: nilai runtun waktu yang stasioner,
$Zt-s$, $Zt-2s$: previous period value,
$\phi 1$, $\phi 2$, ϕp : autoregressive (AR) model coefficients,
$\theta 1$, $\theta 2$, θq : moving average (MA) model coefficients,
αt : MA model parameter error.

Long Short Term Memory (LSTM)

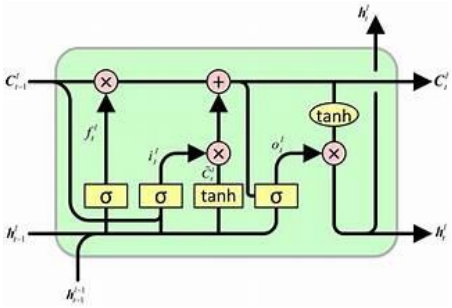

Fig. 1 LSTM Architecture Model

Long Short Term Memory (LSTM) is a development of the Recurrent Neural Network (RNN) architecture, which is a recurrent network that is more effective in learning data patterns by providing memory cells to be able to store information for a long time (Uly, N., Hendry, H., & Iriani, 2023). In this LSTM method, optimization is used to optimize its parameters (Wardianto et al., 2023). In 1997 LSTM was first developed and introduced by Hochreiter and Schmidhuber (Rolangon et al., 2023). LSTM is proposed to be a solution to the formation of vanishing gradients when processing long sequential data. Based on information quoted from (Puteri, 2023) there are several architectures in LSTM, namely:

1. Forget Gate
   Forget Gate (ft) serves to delete information from the cell state. The equation of ft is as follows:

$$f_t = \sigma\,(W_f \times [\,h_t{-}1 \times x_t] + b_f) \qquad (2)$$

   description :
   f_t : forget gate
   σ : sigmoid activation function
   W_f : weights for the forget gate
   x_t: input at time t
   $h_t{-}1$: output from the previous hidden layer
   b_f : bias for the forget gate

2. Input Gate
   Input Gate (it) functions to decide new information that enters the cell state. The equation of it is as follows :

$$i_t = (W_i \times [h_t{-}1 \times x_t\,] + b_i)$$
$$\tilde{C}_t = \tanh(W_C \times [\,h_t{-}1 \times x_t] + b_C) \qquad (3)$$

   Information :
   i_t : input gate
   W_i : weights for the input gate
   b_i : bias for the input gate
   $\tilde{C}_t$: candidate memory
   Tanh : Tanh activation function
   W_C : weights for the candidate memory
   b_C : bias for the candidate memory

3. Cell State
   Cell State (Ct) serves as a critical component that enables the network to understand, store, and manage information over long time sequences. The equation of cell state is as follows:

$$C_t = f_t \times C_{t-1} + i_t \times \tilde{C}_t \qquad (4)$$

* Corresponding author

e- ISSN : 2541-2019
p- ISSN : 2541-044X

Information :
C_t : cell state at time *t*
C_{t-1} : cell state at the previous time step
f_t : forget gate
i_t : input gate

4. Output Gate
Output Gate (*ot*) functions to sort useful information from the cell state current and display it as output. The equation of ot is as follows:

$$o_t = (W_o \times [\, h_t{-}1 \times \ x_t] + b_o) \qquad (5)$$

Information:
o_t : output gate
W_o : weights for the output gate
b_o : bias for the output gate

5. Hidden Layer
Hidden Layer (ht) functions to support understanding and processing of sequence data. The equation of ht is as follows:

$$h_t = o_t \times \tanh(C_t) \qquad (6)$$

Information :
h_t : hidden layer at time *t*

Prophet

Prophet is a prediction or forecasting method developed by Facebook which is available in R and Python (Ridla et al., 2023). Prophet can handle predictions by utilizing historical data and holiday parameters in determining predictions. In time series forecasting, Prophet only uses 2 main data components, namely the time component and the data quantity component denoted by 'ds' and 'y' (Auliya et al., 2023). Prophet is defined as follows:

$$y(t)= g(t) + s(t) + \varepsilon_t \qquad (7)$$

Information:
g(t) : trend model that describes the long-term increase or decrease in the data.
s(t) : Fourier series model that describes how the data is affected by seasonal factors such as the time of year.
ε_t : model that represents the irreducible error term.

Mean Squared Error (MSE)

Mean Squared Error is a method that can be used to evaluate the error level of a forecasting method (Andalia & MOULITA, 2023). MSE calculates the average of the squared differences between the actual value and the forecasted value. The smaller the MSE value obtained, the more accurate the forecast is. The formula for this method is:

$$\text{MSE}= \frac{1}{n} \sum_{t=1}^{n} (At - Ft)^2 \qquad (8)$$

Information :
At : actual demand in period t,
Ft : demand forecast in period t,
n : number of forecast periods involved.

Root Mean Square Error (RMSE)

Root Mean Square Error is a method of estimating squares that measures the average size of the error. RMSE is the square root of the average squared difference between the predicted data results and the actual values (Milniadi & Adiwijaya, 2023). RMSE provides an overview of the extent to which the forecast results differ from the actual values. The formula for this method is:

$$\text{RMSE}= \sqrt{\frac{1}{n} \sum_{t=1}^{n} (At - Ft)^2} \qquad (9)$$

description:
At : actual demand in period t,
Ft : forecast demand in period t

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Mean Absolute Percentage Error (MAPE)

Mean Absolute Percentage Error is a method used to evaluate and measure prediction error in percentage form (Andalia & MOULITA, 2023). MAPE provides an overview of the extent to which the forecasted value differs from the actual value. The MAPE value can be calculated using the following equation:

$$MAPE=\frac{1}{n}\frac{1}{n}\sum_{t=1}^{n}\left|\frac{At-Ft}{At}\right|\times 100\% \qquad (10)$$

description:
At : actual demand in period t,
Ft : forecast demand in period t,
n : number of forecast periods involved.

A forecast is said to be very accurate if it has a MAPE value of less than 10%. The criteria for MAPE accuracy in forecasting can be seen in Table 1 below (Chicco et al., 2021).

Table 1
MAPE Value Range

Range	Description
<10%	Very Good
10%-20%	Good
20%-50%	Passable
>50%	Not Good

METHOD

The method used in this study uses an experimental approach to test several forecasting methods in conducting sales forecasting. The stages in this study are divided into 5 stages, namely business understanding, data understanding, data preparation, and modeling, and evaluation. Figure 2 is the stages of this study:

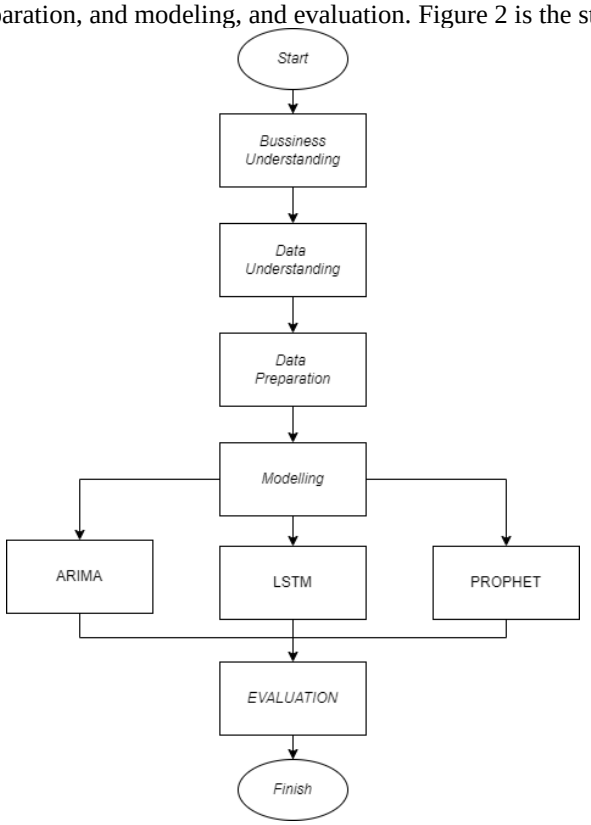

Fig.2 Research Stages

Business Understanding

The first stage carried out in this research is Business Understanding. This stage is a process of understanding the data or problems to be solved. This stage involves understanding the objectives of the problem to be solved, and making research plans to achieve the objectives of the research.

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Data Understanding

In the second stage, the process of collecting data from companies and other sources is carried out. After that, the process of understanding the data used is carried out, and identifying the quality of data that will be used in further analysis. This stage is carried out to ensure that there are no missing values.

Data Preparation

This stage is the stage of data preparation that will be processed in the next stage, such as determining the field, data cleaning and data transformation according to the needs of the next analysis. At this stage, the data is also divided into two parts, namely 80% training data and 20% testing data.

Modeling

The fourth stage is the modeling process. At this stage, model development is carried out and the right algorithm is determined for the problem being analyzed. This process is through model training and model adjustment to obtain a model that has optimal performance.

Evaluation

The fifth stage carried out is evaluation. At this stage, an analysis of the results of the quality of the model developed will be carried out through model validity testing to determine the extent to which the model used can achieve the desired goals. The validity tests used in this forecast are the MAPE, MSE, and RMSE tests.

RESULTS

Business Understanding dan Data Understanding

In this study, a model has been developed to forecast Bandung white bread sales using the ARIMA, LSTM, and Prophet methods. The data used in this study are daily sales data at the Bandung White Bread Factory from March 2021 to May 2023 or 822 days, each day there is one data with different sales amounts.

Data Preparation

The sales data obtained from this bakery has several attributes, namely Sales Date, Total Sales at the Factory, Total Sales at several stores, and total sales. At this stage, a data cleaning process is carried out to remove unnecessary data attributes so that the remaining attributes are only the sales date and total sales. Then this data is divided into two, namely 80% training data and 20% testing data. The sales data used in this study can be seen in Table 2 and the plot of the sales data can be seen in Figure 3.

Table 2. Sales Data

Date	Sales Amount
01-03-2021	718
02-03-2021	719
03-03-2021	837
.....	.....
31-05-2023	758

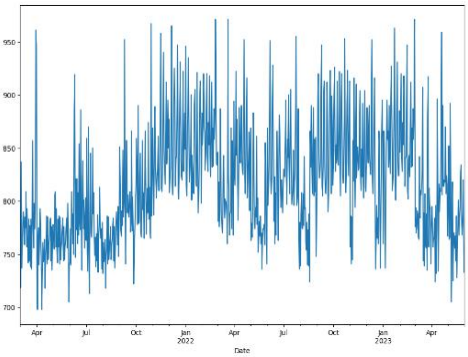

Fig.3 Sales Data Plot

Modeling

Implementation of ARIMA Method

In implementing the ARIMA method in forecasting this factory, the dataset used must be stationary with a significant value (α) of 0.05, if the dataset is not stationary, then a differencing process is needed. The stationary

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

test is carried out through the Augmented Dicky-Fuller (ADF) Test. The results of the Augmented Dicky-Fuller (ADF) Test can be seen in Table 3 and Figure 4 below.

Table 3. ADF Test Results

P-Value	Symbol	Alpha
0.0222454815	<	0.05

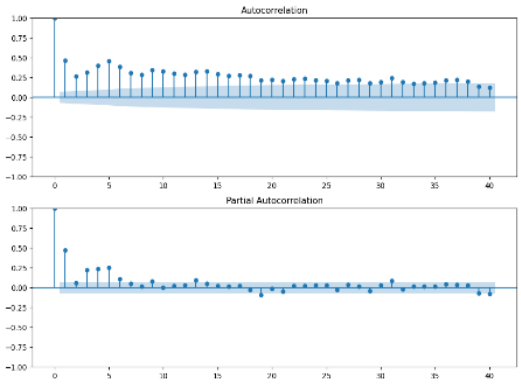

Fig. 4 ACF and PACF plots

Based on the results of the ADF test as seen in Table 3 and Figure 4, it can be concluded that this dataset is stationary because the value generated from the ADF test is 0.0222454815.

The next step is to determine several temporary models that will be used in the forecasting process. Several temporary models obtained can be seen in Table 4.

Table 4. Temporary ARIMA Model

Model ARIMA	MAPE	MSE	RMSE
ARIMA (1,0,1)	4.6124%	2314.9084	48.1134
ARIMA (1,0,2)	4.5484%	2248.0822	47.4139
ARIMA (1,1,1)	5.2650%	5890.0372	76.7465
ARIMA (1,1,2)	5.2467%	5831.8387	76.3664
ARIMA (2,0,1)	4.5725%	2271.9168	47.6646
ARIMA (2,1,1)	5.2364%	5823.6852	76.3130
ARIMA (2,1,2)	5.2361%	5823.6518	76.3128

From the experiments conducted on several temporary models, the best ARIMA model was obtained to be used in the forecasting process, namely ARIMA (1,0,2) with a MAPE value of 4.5484%, an MSE value of 2248.0822, and an RMSE value of 47.4139. Details of the summary generated from the ARIMA (1,0,2) model can be seen in Figure 5.

SARIMAX Results

Dep. Variable:	Jumlah Pax	No. Observations:	164
Model:	ARIMA(1, 0, 2)	Log Likelihood	-865.661
Date:	Sun, 12 Nov 2023	AIC	1741.322
Time:	22:01:38	BIC	1756.821
Sample:	12-19-2022	HQIC	1747.614
	- 05-31-2023		
Covariance Type:	opg		

	coef	std err	z	P>|z|	[0.025	0.975]
const	811.6988	18.879	42.996	0.000	774.697	848.700
ar.L1	0.9643	0.036	26.551	0.000	0.893	1.035
ma.L1	-0.6221	0.085	-7.277	0.000	-0.790	-0.455
ma.L2	-0.2005	0.078	-2.554	0.011	-0.354	-0.047
sigma2	2242.7149	273.342	8.205	0.000	1706.974	2778.456

Ljung-Box (L1) (Q):	0.11	Jarque-Bera (JB):	3.99
Prob(Q):	0.74	Prob(JB):	0.14
Heteroskedasticity (H):	1.04	Skew:	0.37
Prob(H) (two-sided):	0.87	Kurtosis:	2.80

Fig. 5 Summary of ARIMA Model (1,0,2)

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

From Figure 5, it can be seen that the ARIMA (1,0,2) model has a p-value of 0.011, which means that this model is significant.

The forecasting results of the ARIMA method can be seen in Table 5 and a comparison graph of the forecasting results with actual data can be seen in Figure 6.

Table 5. Forecasting Results

Date	Actual Amount	Prediction Result	Difference
19-12-2022	767	812	-45
20-12-2022	796	791	5
21-12-2022	765	801	-36
.....	.....	.....	.....
28-05-2023	788	788	0
29-05-2023	820	793	27
30-05-2023	738	803	-70
31-05-2023	758	774	-16

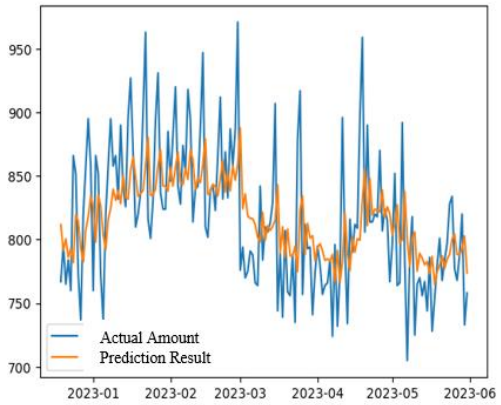

Fig. 6 Comparison of Forecast Results with Actual Data

Figure 6 is a comparison graph of the forecast results with actual data. Based on the forecast results in Table 5, it is known that the lowest difference occurred on March 1, 2023 with a difference value of -112 pax, and the highest difference occurred on April 19, 2023 with a difference value of 128 pax. The average difference between the forecast results and the actual value in the ARIMA method is 1 pax. The accuracy of the forecast results can be seen in Table 6.

Table 6. ARIMA Forecasting Accuracy Results

MAPE	MSE	RMSE
4.548%	2248.0822	47.4139

DISCUSSIONS

Implementation of LSTM Method

In the implementation of the LSTM method in forecasting in this factory, 1 LSTM model was formed with parameter initialization, namely 1 hidden layer, 200 neuron units, and 1 dense layer. The parameters of the LSTM model can be seen in Figure 7.

```
Model: "sequential"
_________________________________________________________________
 Layer (type)                Output Shape              Param #
=================================================================
 lstm (LSTM)                 (None, 200)               161600

 dense (Dense)               (None, 1)                 201

=================================================================
Total params: 161801 (632.04 KB)
Trainable params: 161801 (632.04 KB)
Non-trainable params: 0 (0.00 Byte)
```

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Fig. 7 LSTM parameters

Next, conduct the experiment process with several epochs. The epoch with the smallest MSE and RMSE values will be used to conduct the forecasting process. The results of the experiment on several epoch models can be seen in Table 7.

Table 7. LSTM Forecasting Accuracy Results from Several Epochs

Epoch	MAPE	MSE	RMSE
100	7.4997%	3403.5008	58.3395
200	7.4442%	3356.2904	57.9335
300	7.4934%	3378.5139	58.1250
400	7.3406%	3351.3053	57.8905
500	7.3275%	3246.2740	56.9761

From Table 7, it is obtained that epoch 500 has the smallest MAPE, MSE, and RMSE values, producing a final loss value of 0.0027. The results of the LSTM method forecasting can be seen in Table 8 and a comparison graph of the forecasting results with actual data can be seen in Figure 8.

Table 8. LSTM Forecasting Results

Date	Actual Amount	Prediction Result	Difference
19-12-2022	767	742	25
20-12-2022	796	771	25
21-12-2022	765	798	-33
.....	.....	.....	.....
30-05-2023	733	821	-88
31-05-2023	758	739	19

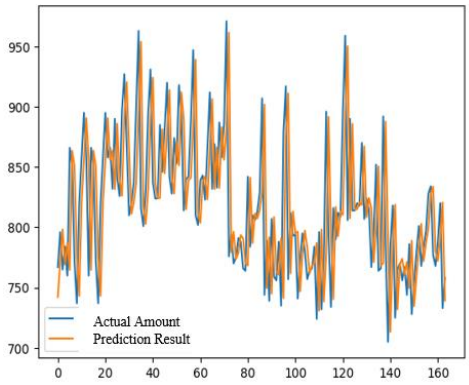

Fig. 8 Comparison Chart of Forecasting Results

Figure 8 is a comparison graph of the forecast results with actual data. Based on the forecast results in Table 8, it is known that the lowest difference occurred on March 1, 2023 with a difference value of -185 pax, and the highest difference occurred on March 24, 2023 with a difference value of 138 pax. The average difference between the forecast results and the actual value in the LSTM method is -1 pax. The accuracy of the forecast results can be seen in Table 9.

Tabel 9
LSTM Forecasting Accuracy Results

MAPE	MSE	RMSE
7.3275%	3246.2740	56.9761

1. Implementation of Prophet Method

In implementing the Prophet method in forecasting this factory, the dataset used in the forecasting process must first have its columns changed to "ds" and "y" according to what is needed by the Prophet algorithm. The results of the Prophet method forecasting can be seen in Table 10 and a comparison graph of the forecasting results with the actual data can be seen in Figure 9.

* Corresponding author

Sinkron : Jurnal dan Penelitian Teknik Informatika
Volume 8, Nomor 4, October 2024
DOI : https://doi.org/10.33395/sinkron.v8i4.14057
e- ISSN : 2541-2019
p- ISSN : 2541-044X

Table 10. Prophet Forecasting Results

Date	Actual Amount	Prediction Result	Difference
19-12-2022	767	761	6
20-12-2022	796	761	35
21-12-2022	765	778	-13
.....	.....	.....	.....
29-05-2023	820	934	-114
30-05-2023	738	933	-200
31-05-2023	758	900	-142

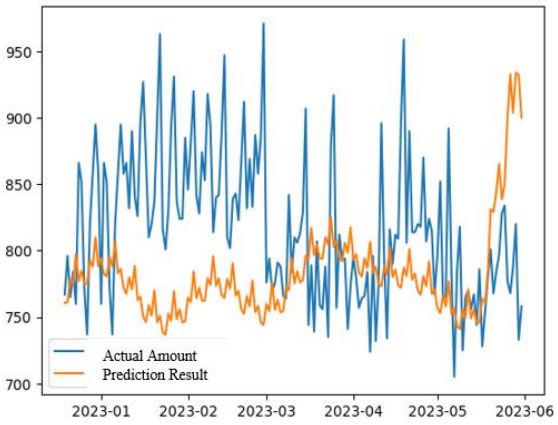

Fig. 9 Comparison Chart of Forecasting Results

Figure 9 is a comparison graph of the forecast results with actual data. Based on the forecast results in Table 10, it is known that the lowest difference occurred on May 30, 2023 with a difference value of -200 pax, and the highest difference occurred on February 28, 2023 with a difference value of 227 pax. The average difference between the forecast results and the actual value in the Prophet method is 36 pax. The accuracy of the forecast results can be seen in Table 11.

Tabel 11. Prophet Forecasting Accuracy Results

MAPE	MSE	RMSE
7.402%	6322.4360	79.5137

Evaluation

Evaluation is the stage of measuring the accuracy of the forecasting results from the ARIMA, LSTM, and Prophet methods. The evaluation in this study uses three tests, namely MAPE, MSE, and RMSE. A comparison of the MAPE, MSE, and RMSE values of the three forecasting methods can be seen in Table 12.

Tabel 12. Comparison Forecasting Accuracy Results

Metode	MAPE	MSE	RMSE
ARIMA	4.548%	2248.0822	47.4139
LSTM	7.3275%	3246.2740	56.9761
Prophet	7.402%	6322.4360	79.5137

Based on Table 12, the results show that in the MAPE, MSE, and RMSE tests carried out on the three methods used in this study, the ARIMA (1,0,2) method is the best method that provides the smallest error accuracy value compared to the LSTM and Prophet methods. The ARIMA (1,0,2) method forecasting produces a MAPE accuracy value of 4.548%, an MSE value of 2248.0822, and an RMSE value of 47.4139. So in this study the ARIMA method has the best performance.

* Corresponding author

CONCLUSION

Based on the results of a comparative study of three methods, namely ARIMA, LSTM, and Prophet in forecasting sales at the Bandung White Bread Factory using 822 sales data, the results showed that the ARIMA (1,0,2) method had the best performance compared to the LSTM and Prophet methods, because the ARIMA (1,0,2) method produced the smallest error accuracy value, namely with a MAPE value of 4.548%, an MSE value of 2248.0822, and an RMSE value of 47.4139. So it can be concluded that the ARIMA method is the most suitable method for use in forecasting sales in the coming period. The suggestion that the author can convey is that in future research it is recommended to create a mobile or web-based forecasting application using the ARIMA (1,0,2) method to overcome the problem of stock shortages at the Bandung White Bread Factory.

REFERENCES

Alim, M. N. (2023). Pemodelan Time Series Data Saham LQ45 dengan Algoritma LSTM, RNN, dan Arima. *PRISMA, Prosiding Seminar Nasional Matematika*, *6*, 694–701.

Andalia, W., & MOULITA, R. A. N. (2023). Peramalan Jumlah Persediaan Komoditas di PT Pelabuhan Indonesia II Cabang Palembang Menggunakan Metode Moving Average dan Exponential Smoothing. *JIeTri: Journal of Industrial Engineering Tridinanti*, *1*(01), 19–26.

Auliya, Y. A., Nurdiansyah, Y., & Astuti, A. P. (2023). Peramalan Jumlah Pengunjung Objek Wisata Gumul Paradise Island Kabupaten Kediri Menggunakan Metode Prophet. *INFORMAL: Informatics Journal*, *8*(1), 37–43. https://doi.org/https://doi.org/10.19184/isj.v8i1.35605

Cherrly, A., Somya, R., Kristen, U., & Wacana, S. (2023). Prediksi Penjualan Tiket Wisata Taman Bermain Menggunakan Metode ARIMA. *Techno.Com*, *22*(2), 312–322. https://doi.org/10.33633/tc.v22i2.7950

Chicco, D., Warrens, M. J., & Jurman, G. (2021). The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation. *PeerJ Computer Science*, *7*, e623. https://doi.org/https://doi.org/10.7717/peerj-cs.623

Ferdinandus, Y. R. M., Kusrini, K., & Hidayat, T. (2023). Gold Price Prediction Using the ARIMA and LSTM Models. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, *8*(3), 1255–1264.

Haryanto, T. (2023). Pemanfaatan Model Long Short Term Memory (LSTM) Untuk Prediksi Harga Emas Sebagai Instrumen Investasi Dalam Mempersiapkan Ancaman Resesi Global 2023. *The Indonesian Journal of Computer Science*, *12*(2). https://doi.org/https://doi.org/10.33022/ijcs.v12i2.3176

Hayadi, B. H., Sudipa, I. G. I., & Windarto, A. P. (2021). Model Peramalan Artificial Neural Network pada Peserta KB Aktif Jalur Pemerintahan menggunakan Artificial Neural Network Back-Propagation. *MATRIK: Jurnal Manajemen, Teknik Informatika Dan Rekayasa Komputer*, *21*(1), 11–20. https://doi.org/https://doi.org/10.30812/matrik.v21i1.1273

Lin, A. K. (2024). The AI Revolution in Financial Services: Emerging Methods for Fraud Detection and Prevention. *Jurnal Galaksi*, *1*(1), 43–51. https://doi.org/10.70103/galaksi.v1i1.5

Makkulawu, A. R. (2023). Studi Kasus Kelayakan Usaha Produksi Roti Coklat Di Ukm XYZ Kabupaten Barru. *Jurnal Indonesia Sosial Teknologi*, *4*(4), 402–412. https://doi.org/https://doi.org/10.59141/jist.v4i4.606

Milniadi, A. D., & Adiwijaya, N. O. (2023). Analisis Perbandingan Model Arima Dan Lstm Dalam Peramalan Harga Penutupan Saham (Studi Kasus: 6 Kriteria Kategori Saham Menurut Peter Lynch). *SIBATIK JOURNAL: Jurnal Ilmiah Bidang Sosial, Ekonomi, Budaya, Teknologi, Dan Pendidikan*, *2*(6), 1683–1692. https://doi.org/https://doi.org/10.54443/sibatik.v2i6.798

Nurarofah, E., Herdiana, R., & Nuris, N. D. (2023). Penerapan Asosiasi Menggunakan Algoritma Fp-Growth Pada Pola Transaksi Penjualan Di Toko Roti. *JATI (Jurnal Mahasiswa Teknik Informatika)*, *7*(1), 353–359. https://doi.org/https://doi.org/10.36040/jati.v7i1.6299

Nurhasanah, D., & Dini, S. K. (2023). Peramalan Jumlah Peserta Kb Aktif Pengguna Alat Kontrasepsi Pil di Daerah Istimewa Yogyakarta Menggunakan Metode ARIMA. *Emerging Statistics and Data Science Journal*, *1*(2), 170–177. https://doi.org/https://doi.org/10.20885/esds.vol1.iss.2.art18

Ridla, M. A., Azise, N., & Rahman, M. (2023). Perbandingan Model Time Series Forecasting Dalam Memprediksi Jumlah Kedatangan Wisatawan Dan Penumpang Airport. *Simkom*, *8*(1), 1–14. https://doi.org/https://doi.org/10.51717/simkom.v8i1.103

Rolangon, A., Weku, A., & Sandag, G. A. (2023). Perbandingan Algoritma LSTM Untuk Analisis Sentimen Pengguna Twitter Terhadap Layanan Rumah Sakit Saat Pandemi Covid-19. *TeIKa*, *13*(01), 31–40. https://doi.org/https://doi.org/10.36342/teika.v13i01.3063

Sudipa, I. G. I., Riana, R., Putra, I. N. T. A., Yanti, C. P., & Aristana, M. D. W. (2023). Trend Forecasting of the Top 3 Indonesian Bank Stocks Using the ARIMA Method. *Sinkron: Jurnal Dan Penelitian Teknik Informatika*, *8*(3), 1883–1893. https://doi.org/10.33395/sinkron.v8i3.12773

Suryadana, K., & Sarasvananda, I. B. G. (2024). Streamlining Inventory Forecasting with Weighted Moving Average Method at Parta Trading Companies. *Jurnal Galaksi*, *1*(1), 12–21.

* Corresponding author

e- ISSN : 2541-2019
p- ISSN : 2541-044X

https://doi.org/10.70103/galaksi.v1i1.2

Suryawan, I. G. T., Arimbawa, I. K. S., & Sudipa, I. G. I. (2023). Implementation of Naive Bayes Method for Granting Fisherman Business Credit. *Jurnal Info Sains: Informatika Dan Sains*, *13*(01), 24–32. https://doi.org/10.54209/infosains.v13i01

Uly, N., Hendry, H., & Iriani, A. (2023). CNN-RNN Hybrid Model for Diagnosis of COVID-19 on X-Ray Imagery: Hybrid Model CNN-RNN untuk Diagnosis COVID-19 pada Citra. *Digital Zone: Jurnal Teknologi Informasi Dan Komunikasi*, *14*(1), 57–67. https://doi.org/https://doi.org/10.31849/digitalzone.v14i1.13668

Wahyuni, I. D., Yuniarti, T., & Rapi, A. (2022). Penerapan Model ARIMA Dalam Memprediksi Penjualan Produk Minuman Teh Botol Sosro Ukuran 350 mL. *INVENTORY| Industrial Vocational E-Journal On Agroindustry*, *3*(2), 69–82.

Wardianto, W., Farikhin, F., & Kusumo Nugraheni, D. M. (2023). Analisis Sentimen Berbasis Aspek Ulasan Pelanggan Restoran Menggunakan LSTM Dengan Adam Optimizer. *JOINTECS (Journal of Information Technology and Computer Science)*, *8*(2), 67. https://doi.org/10.31328/jointecs.v8i2.4737

Yang, X. (2019). The prediction of gold price using ARIMA model. *2nd International Conference on Social Science, Public Health and Education (SSPHE 2018)*, 273–276.

Zuhri, S., & Nisa, R. (2022). Peramalan Volume Sampah Menggunakan Pendekatan Arima Time Series. *The Journal of Science and Technology (IsAT)*, *4*(1), 14–19.

* Corresponding author