Analysis Of Bread Demand Forecasting Using Recurrent Neural Network (RNN) Method Based On Operational Delivery Data

Harinudin Saputro[1], **Mohammad Zoqi Sarwani***[2], **Rudi Hariyanto**[3]
[1,2,3] Informatics Engineering, Faculty of Information Technology, Universitas Merdeka Pasuruan
E-mail : harinudinsaputro@student.unmerpas.ac.id[1], zoqisarwani@unmerpas.ac.id[2], rudihariy4nt0@gmail.com[3]*
**Corresponding author*
Received 6 July 2025; Revised 25 July 2025; Accepted 28 July 2025

Abstract-Accurate demand forecasting plays a vital role in optimizing inventory and distribution planning, especially for perishable goods such as bread. This study develops a time series forecasting model using a Recurrent Neural Network (RNN) with a Sequential architecture to predict daily bread demand. Unlike previous research, this model is trained on two years of real operational delivery data (2023–2024), enabling it to capture actual consumption patterns more effectively. The model leverages a 7-day sequence window to predict the next day's demand, reflecting weekly seasonality. Data preprocessing includes normalization and cleaning, followed by training with the Stochastic Gradient Descent (SGD) optimizer. The model achieved a Mean Absolute Percentage Error (MAPE) of 4.88% and an accuracy of 86.90%, demonstrating high predictive performance and robustness in handling fluctuating, real-world data. The implementation of this model provides a practical solution for improving production planning, reducing waste, and enhancing supply chain responsiveness. The findings confirm that RNN-based models are effective tools for demand forecasting in dynamic business environments.

Keywords - Forecasting, Recurrent Neural Network (RNN), Demand Prediction, Operational Delivery Data, Bread Industry.

1. INTRODUCTION

The context of intensifying business competition, the inability to accurately forecast product demand poses substantial challenges for companies, particularly within the manufacturing and distribution sectors. Moreover, the principles of demand forecasting assisted in obtaining essential information that was employed to enhance understanding of forecasting, reinforcing and connecting that theory to the research through a deductive method [1]. Forecasting demand has always been critical to the operational strategy for managing stock and improving the level of customer service satisfaction [2]. In this context, the ability to accurately forecast stock requirements become critically important, particularly when demand data fluctuates over time. Forecasting thus serves as a strategic approach to estimate the quantity of consumer product demand, helping to prevent both shortages and overproduction. Product demand forecasting is the process of predicting future product needs to ensure smooth business operations and minimize the risk of overstocking or stockouts. Predictive analytics, utilizing machine learning and vast data resources, has surfaced as a revolutionary instrument in predicting business and consumer behaviors [3].

With the rapid advancement of technology, Artificial Intelligence methods have become increasingly prevalent in data analysis, offering powerful tools for uncovering patterns, improving prediction accuracy, and supporting data-driven decision making. One of the approaches proven to be effective in handling sequential data is the RNN [4]. RNN is an artificial neural network model designed to recognize patterns in historical data and generate predictions based on long-term dependencies in time series data. Numerous companies and sectors today depend on vast

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

amounts of time series data, making time series forecasting a crucial field of study [5]. RNN are Artificial Neural Network (ANN) designed for tasks involving sequence forecasting [6]. RNN is particularly well-suited for processing historical data, especially when temporal patterns and the sequence of events significantly influence the desired outcomes. To enable the systems to operate autonomously and intelligently, we created a RNN that leverages past data to forecast future indoor PM2.5 levels [7].

RNN were introduced in the 1980s as a specialized type of neural network designed for modeling sequential or time series data [8, 9]. Unlike traditional feedforward neural networks, RNN are equipped with recurrent connections that allow information from previous time steps to be retained and reused in the current computation. This is achieved by connecting the hidden layer to itself through a time delay, enabling the network to capture temporal dependencies across time steps. This architecture makes RNN particularly suitable for tasks where past context plays a critical role in predicting future outcomes [9].

The use of RNN in forecasting is becoming increasingly widespread due to their ability to handle both linear and nonlinear relationships in sequential data. The core of the Long Short-Term Memory (LSTM) algorithm lies in its ability to integrate nonlinear control and dependencies within the RNN cell [10]. We selected a deep RNN with 3 layers for all experiments in the extremely short-term prediction [11]. Therefore, the role of RNN in this study is crucial for recognizing patterns and generating accurate forecasts to optimize decision-making processes. Compared to conventional forecasting methods such as ARIMA (Autoregressive Integrated Moving Average) or Linear Regression, RNN offers advantages in dealing with complex and nonlinear data, making it more effective for forecasting product demand.

Another widely used conventional forecasting method in demand planning is the ARIMA model. ARIMA utilizes past and present values of the dependent variable to generate accurate short-term predictions. A commonly utilized and popular econometric model for predicting economic outcomes the ARIMA model analyzes events and processes through time series [12]. This method has advantages in its ability to identify and model trends well but also has some limitations regarding parameter selection and interpretation of results [13].

Conversely, linear regression methods are also used for demand forecasting by considering independent variables that influence demand. The findings indicate that while ARIMA outperformed regression on both daily and weekly datasets using 1-year data, linear regression surpassed ARIMA in performance when 2-year data was used [14]. In this context, the use of RNN offers a more effective solution for time series forecasting. One of the widely adopted techniques is the use of RNN for time series prediction [15]. Time series forecasting, which refers to predicting future values in a data series, facilitates data-driven decision strategies [16].

This study presents several key novelties that distinguish it from prior research in the field of demand forecasting. First, the data used in this study is derived from actual operational delivery records, rather than estimated sales or inventory data. This provides a more accurate and realistic representation of real-world distribution volume, allowing the forecast to be directly applied in planning stock requirements for the upcoming year. Second, the model is developed using a RNN with a sequential architecture, RNN are designed for processing sequences [17]. This approach enables the model to learn from patterns over the previous week to predict the following day's demand an effective strategy for capturing weekly dynamics in daily data. Third, the model achieved a high level of forecasting accuracy, indicating strong performance even on fluctuating operational data. The ability of the federated RNN model to achieve high accuracy comparable to that of a centralized model. The experimental findings indicated that the suggested deep RNN-based short-term forecasting algorithm attained greater prediction accuracy [11].

These contributions highlight the novelty of this research not only in its theoretical approach to time series forecasting but also in its practical applicability for inventory planning and distribution logistics, particularly for perishable goods such as bread. Accordingly, this study aims to develop a demand prediction model based on RNN using operational data from bread

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

delivery. By integrating this model into the demand planning system, the company is expected to improve its responsiveness to demand fluctuations, enhance inventory efficiency, and reduce distribution costs.

2. RESEARCH METHOD

This study involves several stages to address the identified problems. The stages include a research flow, literature review, problem identification, data collection, data analysis, modeling, model evaluation, and conclusion.

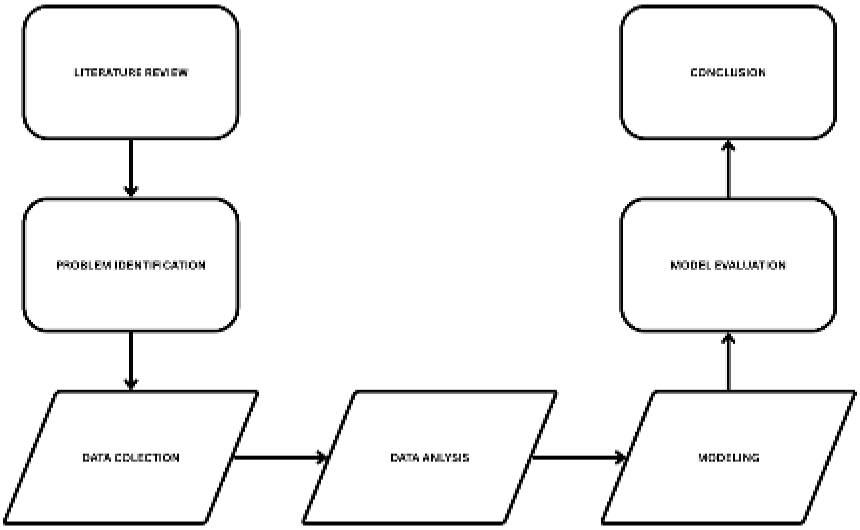

Figure 1. Research Flow

The literature review stage involves searching for and analyzing references with the aim of understanding the fundamental concepts, techniques, and methodologies to be used, as well as identifying research gaps that can be addressed. Problem identification is the stage aimed at recognizing the main issues to be solved through the research. Data collection refers to the process of gathering raw data relevant to the study, which will then be used to determine appropriate parameters for the data analysis stage. Data analysis involves processing and interpreting the data to identify patterns, trends, or specific relationships. Modeling is the process of designing and constructing a model that can be used to solve the research problem. Model evaluation is the process of measuring the model's performance using testing data. Finally, the conclusion stage presents a summary of the research results and key findings.

2.1 Literature Review

This stage involves conducting research and gathering reference information from journals, books, and scientific articles related to forecasting using Recurrent Neural Network (RNN). The focus of the review is on time series forecasting techniques, model evaluation, as well as the advantages and challenges of RNN compared to other forecasting methods.

2.2 Problem Identification

In this study, the problem identification stage aims to obtain a detailed understanding of the issues to be addressed. The primary problem lies in the inaccuracy of demand forecasting, which results in both overstock and understock conditions. These issues may lead to excessive inventory waste due to expired products and delivery delays caused by insufficient bread production at certain times. Based on the company's historical data, there is a need for

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

a modeling approach capable of capturing time series patterns to improve forecasting accuracy.

2.3 Data Collection

Table 1. Raw Data

BILLING DATE	*CUSTOMER CODE*	*Customer NAME*	*……*	*INV TYPE*
01.01.2023	30000053	DC SAT MALANG	……	ZSTD
02.01.2023	30000053	DC SAT MALANG	……	ZSTD
03.01.2023	30000053	DC SAT MALANG	……	ZSTD
04.01.2023	30000053	DC SAT MALANG	……	ZSTD
…	…….	…….	……	……
31.12.2024	20004475	SHELL DIPONEGORO SIDOARJO	…….	ZSTD

The data used in this study was obtained from a nearby bread company. Communication was established with relevant personnel to understand the procedures and requirements for data access, as well as to identify the types of data permissible for research purposes. A formal request was submitted to the company to gain access to the data for this study. The data acquired consists of time series records of product delivery operations to several distributors, covering the period from January 2023 to December 2024.

Data collection in this study was carried out by gathering operational delivery records of bread products from the company to various distributors, including supermarkets and minimarkets such as Alfamart, Indomaret, Superindo, and other conventional retail partners. The collected data includes information such as delivery time, depot name, depot code, depot type, distributor type, bread code, bread name, quantity (QTY), net value, and invoice number over a specific period.

This process aims to provide an accurate and comprehensive database as a foundation for the analysis and development of the product demand forecasting model. The data was obtained from the company's integrated operational recording system, allowing for both real-time and historical representation of distribution patterns.

2.4 Data analysis

Data analysis involves processing the data to determine the appropriate parameters for use in model development. A critical aspect of this process is data cleaning, which includes checking for missing values, identifying outliers, and removing irrelevant columns. The purpose of data cleaning is to ensure accurate parameter selection, streamline model processing, and reduce computational time by eliminating unnecessary variables.

For this study, the columns used for modeling include delivery information and the quantity of products delivered on each specific date (QTY). These columns will be aggregated to generate a single column representing the total quantity of bread produced per day.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

Table 1. Data from Analysis Results

Date	*Qty*
01.01.2023	242277
02.01.2023	274635
03.01.2023	246616
04.01.2023	249848
05.01.2023	227158
…	…….
31.12.2024	268001

2.5 Modeling

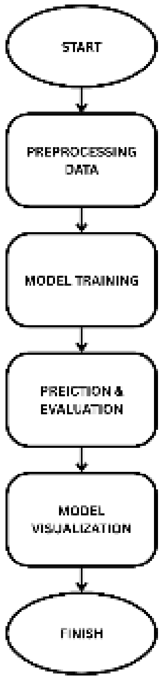

Figure 2. Modeling Flow

1) *Preprocessing Data*
   a. The data is collected and cleaned to remove missing values and anomalies.
   b. Data normalization is applied to scale the values within a certain range, such as 0 to 1, in order to accelerate model convergence.
   c. The dataset is split into training, validation, and test sets with a ratio of 80:10:10.
2) *Model Training*
   a. Model structure
      1) The model is trained using the Number of Input Neurons:
         The model uses a single input neuron, representing one input variable at each time step.
      2) Sequence Length (Window Size):
         The model considers the previous 7 time steps (sequence length = 7) to capture temporal patterns in the data over a given period.
      3) Number of Hidden Layers and Neurons:

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

The model consists of multiple hidden layers, with the number of neurons in each layer varying depending on the results of hyperparameter tuning experiments.

According to [18], the basic formulas of the RNN are as follows:

- *Hidden State Formula*:

$$h_t = \tanh(W_{xt} \cdot x_t + W_{hh} \cdot h_{t\,-\,1} + b_h) \quad (1)$$

Explanation :

1) *ht*: Current hidden state
2) *xt*: Current input
3) *ht−1*: Previous hidden state (memory)
4) *Wxh*: Weight matrix from input to hidden layer
5) *Whh*: Weight matrix from previous hidden to current hidden layer
6) *bh*: Bias for the hidden layer
7) *tanh*: Activation function (Rectified Linear Unit (ReLU) is sometimes used as an alternative)

- *Output layer Formulas*:

$$y_t = W_{hy} \cdot h_t + b_y \quad (2)$$

Explanation :

1) *yt:* Current output
2) *Why:* Weight matrix from hidden layer to output
3) *ht*: Current hidden state
4) *by:* Bias for the output layer

4) Activation Function for Hidden Layers:
   Each hidden layer uses the ReLU activation function to accelerate the learning process and address the vanishing gradient problem.
5) Number of Output Neurons:
   The model uses a single output neuron, which represents the predicted value at the next time step.
6) Activation Function for the Output Layer:
   The output layer uses the ReLU activation function, as the predicted values are expected to be positive.

b. Training data with the Stochastic Gradient Descent (SGD) optimization algorithm.
c. Hyperparameter tuning is performed to obtain the optimal configuration.
d. The model is evaluated on the validation set to prevent overfitting.

3) Prediction and Evaluation

After training, the model is tested on the test set to assess accuracy and performance using evaluation metrics such as MAPE.

4) Model Visualization

The prediction results are visualized in graphical form to compare actual values with predicted values.

2.6 Model Evaluation

Model evaluation is conducted to assess the prediction performance using the following metric of MAPE:

Used to measure the prediction error in percentage terms, MAPE indicates the average absolute difference between predicted and actual values relative to the actual values.

MAPE Formulas :

$$MAPE = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{yi - \hat{y}i}{yi} \right| \quad (3)$$

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

Explanation :
1) *MAPE* : The average absolute percentage error between the actual and predicted values.
2) *n:* Total number of data samples.
3) *yi*: Actual value of the *i* data point.
4) *ŷi:* Predicted value of the *i* data point.

3. RESULTS AND DISCUSSION

The implementation of the Recurrent Neural Network (RNN) model using the Sequential architecture in TensorFlow/Keras yielded highly satisfactory results in forecasting bread demand. The model was trained using daily delivery data from January 2023 to December 2024, and the prediction was extended to cover the entire year of 2025.

Table 3. Data from Predictions Results

Date	*Qty*
01.01.2025	216425
02.01.2025	249568
03.01.2025	255841
04.01.2025	280667
05.01.2025	195376
…	…….
31.12.2025	268001

The RNN model was built using the Sequential API, which allows for a linear stack of layers to be constructed in a forward pass manner. The architecture consisted of the following components:

- *Input Layer:* Accepts a single feature (daily quantity) over a time window of 7 days (sequence length).
- *Hidden Layer:* A SimpleRNN layer with 32 units and ReLU activation function, responsible for capturing temporal dependencies from the sequential data.
- *Output Layer:* A single neuron with ReLU activation to output the predicted quantity for the next time step.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

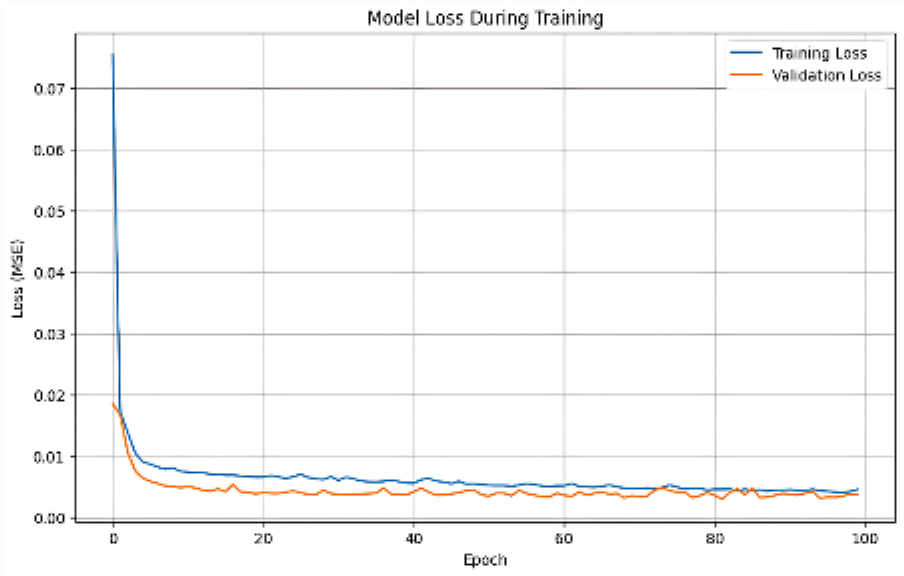

Figure 3. Loss Function Graph

The model was compiled using the SGD optimizer with a learning rate of 0.01 and trained over 100 epochs. The loss function used was Mean Squared Error (MSE), which is suitable for regression-based forecasting. Additionally, Figure 3 illustrates the training and validation loss curves over 100 epochs, which show smooth convergence without significant overfitting, indicating that the model generalized well to unseen data.

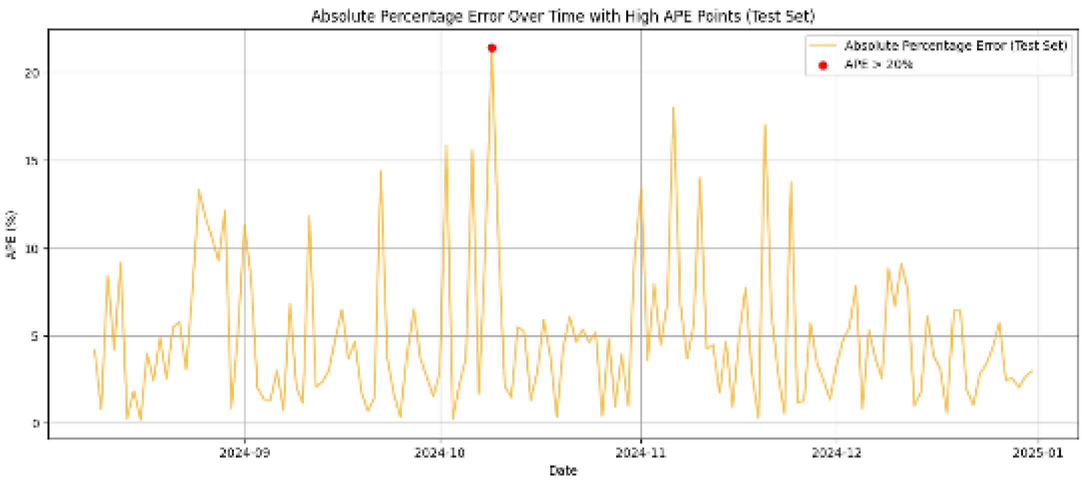

Figure 4. APE (Absolute Percentage Error)

Furthermore, Figure 4 illustrates the daily Absolute Percentage Error (APE) observed during testing. Although several spikes exceeded 20%, most error values remained under 10%, confirming that the model’s predictive performance is consistent and reliable across time.

After training, the model achieved a MAPE of 4.88% and the models achieved an accuracy of 86,90%. This value demonstrates strong predictive performance. In practical terms, the MAPE score implies that the average forecasting error is below 5% relative to actual values, which is considered excellent in demand forecasting.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

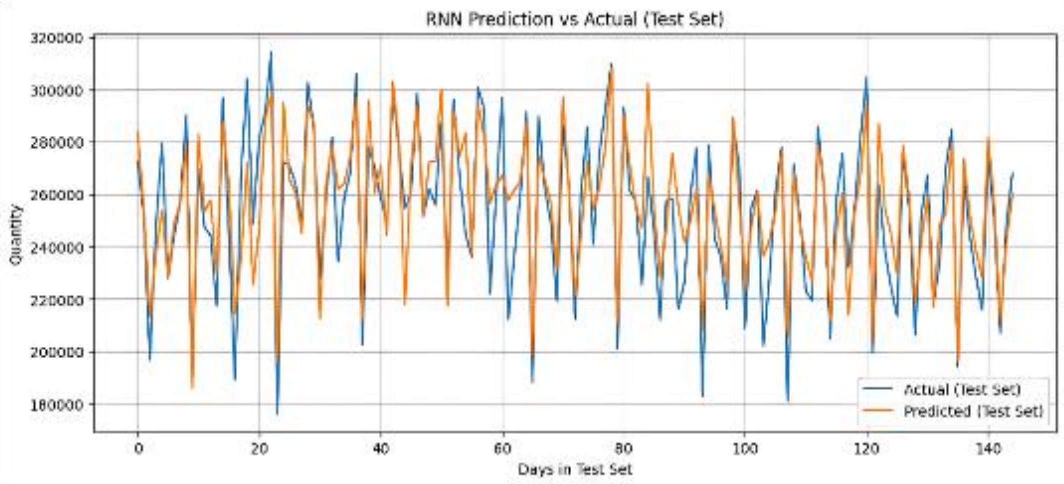

Figure 5. Model Performance of Actual vs Test Set

Figure 5 displays a comparison between the actual values (blue line) and the predicted values (orange line) over the test set period, using a RNN-based forecasting model. The X-axis represents the number of days in the test set, while the Y-axis shows the quantity.

As observed from the figure, the predicted values closely follow the actual values across the entire test period. The RNN model is able to capture the fluctuations and patterns present in the actual data with high fidelity, demonstrating its ability to learn the temporal dynamics inherent in the time series.

This close alignment between predicted and actual data indicates that the model performs well in forecasting short-term demand. The model's predictive accuracy is further supported by the calculated evaluation metric, where it achieved a MAPE of 4.88%, confirming the model's robustness and reliability for operational forecasting tasks.

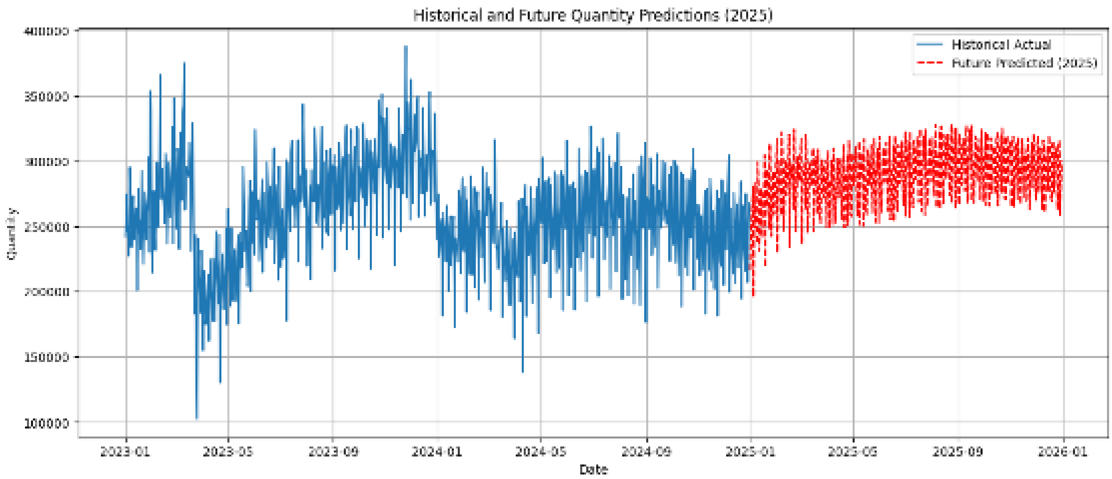

Figure 6. Future Prediction

The figure above presents a visualization of quantity predictions based on historical data and the forecasting model employed. The blue line represents the actual historical data from 2023 to 2024, while the red dashed line indicates the predicted values for the year 2025. From the graph, it is evident that the model successfully captures the seasonal patterns observed in the historical

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

data, with consistent fluctuations over time. Notably, the prediction results show a decline in quantity around the fourth month (April 2025). This downward trend aligns with a similar pattern found in the previous historical data, suggesting that the model is capable of recognizing and following cyclical or seasonal trends accurately.

The Sequential RNN model demonstrated robustness in handling daily operational data, which is typically noisy and irregular. Its ability to retain information from previous time steps (via hidden states) makes it suitable for real-time forecasting tasks in supply chain and logistics settings.

These results validate the effectiveness of RNN-based models in practical business scenarios, where precise demand forecasting can drive more efficient inventory management, reduce waste, and enhance distribution planning.

4. CONCLUSION

This study successfully developed a time series forecasting model for daily bread demand using a RNN implemented through a Sequential architecture. The model was trained on three years of operational delivery data (2023–2024) and demonstrated strong predictive capability, achieving a MAPE of 4.88% and archived an accuracy 86,90%.

The structure of the model comprising a SimpleRNN layer and a fully connected output layer proved effective in capturing temporal patterns and fluctuations in the dataset. The model's ability to produce low error across a wide range of dates, as well as its performance stability shown in the APE analysis, further confirms the model's robustness in operational forecasting.

By integrating this forecasting model into the company's operational planning, decision-makers can reduce the risk of overproduction, minimize delivery inefficiencies, and respond more effectively to dynamic market conditions. Future work may involve testing more complex deep learning architectures such as LSTM or Gated Recurrent Unit (GRU), and incorporating external fa ctors like weather conditions, holidays, or marketing campaigns to improve forecasting accuracy even further.

REFERENCES

[1] D. Keleperas, "Exploring The Iimportance of Demand Forecasting in B2B," *THESEUS,* 2023.

[2] A. Mitra, A. Jain, A. Kishore dan P. Kumar, "A Comparative Study of Demand Forecasting Models for a Multi-Channel Retail Company: A Novel Hybrid Machine Learning Approach," *Operations Research Forum,* vol. 3, p. 58, 2022.

[3] P. waran dan M. Sriramkumar, "Predictive Analytics: Utilizing Machine Learning and Big Data for Forecasting Future Trends in Business and Consumer Behavior," *Strategic Brand Management in the Age of AI and Disruption,* 2025.

[4] P. Norvig dan S. Russell, " Artificial Intelligence: A Modern Approach," 2020.

[5] R. Godahewaa, C. Bergmeir, G. I. Webb, R. J. Hyndman dan P. Montero-Manso, "Monash Time Series Forecasting Archive," *ARXIV,* 2021.

[6] M. Waqas dan U. W. Humphries, "A critical review of RNN and LSTM variants in hydrological time series predictions," *ELSEVIER,* 2024.

[7] X. Dai, J. Liu dan Y. Li, "A recurrent neural network using historical data to predict time series indoor PM2.5 concentrations for residential buildings," *INDOOR AIR: International Journal of Environtment and Health,* 2021.

[8] E. Jeffry L, "Finding Structure In Time," *Cogn. Sci.,* 1990.

[9] S. Mahmood Al-Selwi, M. Fadzil Hassan, S. Jadid Abdulkadir, A. Muneer, E. Hamid Sumiea, A. Alqushibi dan M. Gamal Ragab, "RNN-LSTM: From aplications to medling

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507

techniques and beyond-Systematic review," *Journal of King Saud University - Computer and Information Sciences,* 2024.

[10] L. Gao, Z. Guo, H. Zhang, X. Xu dan H. T. Shen, "Video Captioning with Attention-based LSTM and Semantic Consistency," 2017.

[11] H. K. Ahn dan N. Park, "Deep RNN-Based Photovoltaic Power Short-Term Forecast Using Power IoT Sensors," *Energies,* 2021.

[12] S. Toshaliyeva, "Using the Arima model to forecast the share of railways in the industry," *E3S Web of Conferences 531,* 2024.

[13] S. Arifin, M. M. Manurung, S. Jonathan, M. Effendi dan P. W. Prasetyo, "Trend Analysis of the ARIMA Method: A Survey of Scholarly Works," *Recent in Engineering Science and Technology,* 2024.

[14] M. Beniwal, A. Singh dan N. Kumar, "Predicting Long Term Prices of Nifty Index using Linear Regression and ARIMA: A comparative study," *International Journal of Accounting : Business and Finance,* 2022.

[15] R. Merdiansah, K. Wulandari, M. Hasibuan dan Y. Umaidah, "Perbandingan Kinerja Model RNN, LSTM, dan BLSTM dalam Memprediksi Jumlah Gempa Bulanan di Indonesia," *Jurnal Penelitian Rumpun Ilmu Teknik (JUPRIT),* 2024.

[16] A. Casolao, V. Capone, G. Iannuzo dan F. Camastra, "Deep Learning for Time Series Forecasting: Advances and Open Problems," *MDPI Canada,* 2023.

[17] S. Ashraf Zargar, "Introduction to Sequence Learning Models: RNN, LSTM, GRU," *Departement of Mechanical and Aerospace engineering,* 2021.

[18] R. D. Putra, R. Ilyas dan F. Kasyidi, Pembangkit Kalimat Ilmiah Menggunakan Reccurent Neural Network, SisInfo, 2021.

e-ISSN : 2356-2579 | p-ISSN : 1412-2693, DOI :10.62411/tc.v24i3.13507