# Applying the Long Short-Term Memory Technique to Sales Forecasting for an Auto Parts Vendor

Napat Srinual[1*], Sittichok Sinrat[2] and Chanisara Thitkrathok[3]

[1] Department of Logistics and Transportation Management, Panyapiwat Institute of Management
[2,3] Department of Logistics and Supply Chain Management, Huachiew Chalermprakiet University
*Email: napatsri@pim.ac.th
Received: July 3, 2025; Revised: October 2, 2025; Accepted: November 13, 2025

## Abstract

In the face of increasingly fluctuating car demand, sales orders for automotive parts from various vendors have highlighted the growing importance of sales forecasting and predictive analytics to prevent overstocking and stock outs. However, most of the existing research continues to rely predominantly on traditional time-series forecasting methods that do not fully capture dynamic demand. To address these issues, this study applies a deep learning time-series forecasting technique to reduce the discrepancy and find a suitable method for forecasting the orders. The objective of this study is to propose an algorithm for forecasting sales based on an LSTM network, using historical sales data from an auto parts vendor. This study was conducted in an automotive vendor located in Samut Prakan, utilising a dataset covering 130 weeks from 2023-2025 to. A total of 70 percent of the sales data were employed for training, whereas 30 percent were allocated to the testing sets. The selected item stems from high-value auto parts consisting of 77443582-441V, 77443582-521V, 55183973-060V, and 55183973-440V. The four auto parts were selected based on their high sales volume and strategic importance to the vendor’s inventory management. These items contribute to over 40% of total sales revenue and are frequently subject to stockouts, making accurate forecasting critical. Compared with actual sales orders, the results show that the MAPE forecast error from the LSTM model is lower than the 6.14% error, and it achieves an R-squared of 90.228%, outperforming the existing forecasting technique used by the purchasing department. These findings suggest that the proposed forecasting model can effectively improve the sales order accuracy of automotive part vendors.

**Keywords:** Auto parts, Sales forecasting, Forecast accuracy, Long Short-Term Memory

## 1. Introduction

Predicting a sales order for auto parts is the key research for auto parts vendors [1]. It estimates the actual sales of specific point in advance according to needs of manufacturers.

The main issue with a company name “Vendor ABC auto parts” is inaccurate sales forecasting, resulting in a mismatch between the anticipated demand from manufacturers and the actual sales orders. Consequently, the company experiences overstocking and understocking. The company's existing forecasting method has low accuracy and relies on A 3-week simple moving average of demand (SMA-3weeks) technique. The company currently uses a 3-week SMA based on internal evaluation showing that this period provides the most stable forecast for their replenishment cycle, which operates on a weekly ordering schedule with a lead time of 7 days. This approach has proven to be inadequate for accurately predicting automobile orders, resulting in the accumulation of unsold goods.

Table 1 lists the forecast errors for the items selected in this study. Given these challenges, the company is now exploring advanced deep-learning techniques to improve sales forecasting. Since sales orders are time series, they can be treated as sequential data, making them well suited for advanced time series forecasting methods.

In addition, Abbasimehr et al. [2] pointed out that while advanced deep learning techniques, such as artificial neural networks (ANN) and recurrent neural networks (RNN), offer promising solutions for sales forecasting, they have certain limitations. To overcome these drawbacks, a Long Short-Term Memory (LSTM) network, a specialised type of RNN, was used [3]. The objective of this study is to propose the application of an LSTM network for predicting the sales patterns of automotive parts. Additionally, this study aims to assess the effectiveness of LSTM models in sales forecasting by comparing their performance with of the three-week simple moving average forecasting method currently used by the case study company.

**Table 1** SMA (n=3weeks) forecasting errors for four auto parts

| Parts code | MAPE (%) |
|---|---|
| 77443582-441V | 9.37 |
| 77443582-521V | 20.65 |
| 55183973-060V | 21.10 |
| 55183973-440V | 6.69 |

## 2. Scope of Research

This research scope focuses on analysing and forecasting the sales orders of four different items of automotive parts using time series methods, specifically employing Long Short-Term Memory (LSTM) networks. The data utilized for this analysis consists of monthly procurement records of automotive components collected over a three-year period, from 2023 to 2025. The dataset will be divided into a training set covering 90 weeks and a testing set for 40 weeks. This division allows for robust model training and validation, ensuring the effectiveness of the forecasting methodology. The findings aim to provide insights into purchasing trends and enhance inventory management strategies in the automotive sector.

## 3. Literature Review

### 3.1 Time Series Forecasting

Time series methods are based on historical data of an event and forecasting [4]. Lim and Zohren [5] noted that time-series forecasting has been performed under the assumption of linearity and statistical methods with traditional approaches such as simple moving average (SMA), weighted moving average (WMA), exponential smoothing (ES), single exponential smoothing (SES), Holt linear (HL), and Holt-Winters (HW). However, it is not suitable if the time series contains nonlinear patterns that can be captured by nonlinear models, such as ANNs [6].

### 3.2 Artificial Neural Network (ANN)

An Artificial Neural Network is a machine learning algorithm based on biological neural networks [6]. It is composed of numerous highly interconnected processing elements (neurons) that work together to solve specific problems [7]. There are three main components of ANNs: neurons, interconnections, and learning rules. It can implicitly detect complex nonlinear relationships between dependent and independent variables and is best suited for complex information processing.

### 3.3 Recurrent Neural Network (RNN)

RNN belong to the family of artificial neural networks with recurrent architectures. Recurrence refers to the use of previous outputs as inputs so that they are recurrent. RNN are sequential models that process one or more sequential inputs of variables [8]. Although RNNs are suitable for time-series forecasting, Hurtado-Mora et al. [9] stated that they suffer from the vanishing and exploding gradient problem which makes them difficult to train. To overcome this problem, an extended variant of RNNs called LSTM has been introduced that has gained popularity owing to its excellent performance in time-series forecasting [10].

**3.4 Long Short-Term Memory (LSTM) Network**

LSTM is an RNN that can manage both long- and short-term dependencies. As previously mentioned, the major problem with RNN is the vanishing gradient. LSTMs solve this problem by introducing memory cells and gating mechanisms [11]. Each memory cell $C_t$ holds the temporal states of the network and has three gates, $I_t$, $F_t$, and $O_t$, which control the flow of information.

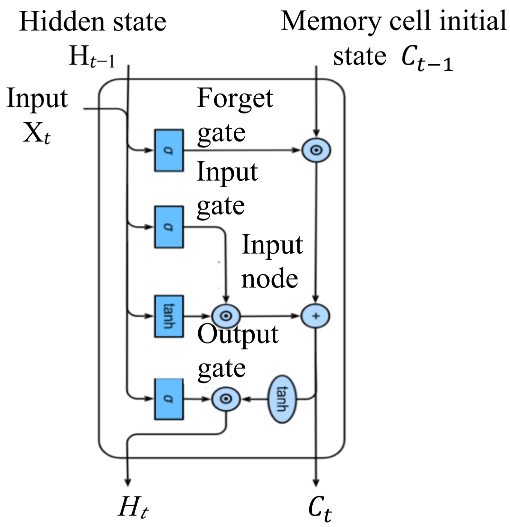


**Figure 1** The structure of LSTM architecture (adapted from [12])

The architecture of LSTM is shown in Figure 1. The gate operations are as follows:

1) Forget gate

This gate decides whether to retain the information in the cell state or discard it. The decision to maintain is based on the input data and the output from the previous node, passed through a sigmoid function, as shown in equation (1).

$$F_t=\sigma\ (W_F\cdot[H_{t-1},H_t]+b_F) \quad (1)$$

2) Input gate

This gate receives new input data and "writes" into each node. This operation can be divided into two parts. First, if the cell state must be updated, the sigmoid function acts as a control mechanism and activates the input gate to decide whether to update the cell state. Second, if the input gate decides to update the cell state, the "tanh" function will generate candidate values () for the state, as shown in equations (2), (3) and (4).

$$I_t=\sigma\ (W_I\cdot[H_{t-1},X_t]+b_I) \quad (2)$$

$$\widetilde{C}_t=\tanh\ (W_C\cdot[H_{t-1},X_t]+b_C) \quad (3)$$

$$C_t=F_t*C_{t-1}+I_t*\widetilde{C}_t \quad (4)$$

3) Output gate

This gate determines what goes out of the cell state. It is a gate that is preparing to output data. The sigmoid determines the parts of the cell state to output. The cell state is passed into the sigmoid to decide what to output as one or zero. Then, the result from the "tanh" is calculated to determine if the output is 1 or -1. Finally, the output from the sigmoid gate is combined with the result from the "tanh" to get the output as shown in equation (5).

$$O_t=\sigma\ (W_O\cdot[H_{t-1},X_t]+b_O) \quad (5)$$

**3.5 Forecasting model metrics**

In the study, the forecasting model performance is evaluated using two selected metrics. Among the various metrics available, this work focuses specifically on the Mean Absolute Percentage Error (MAPE) and the coefficient of determination (R-squared) to assess the accuracy and predictive capability of the forecasting models. These two metrics are used to compare the models and identify the most effective one for sales forecasting, as detailed below.

1) Mean Absolute Percentage Error (MAPE) is to evaluate a performance metric in forecasting models. This can be defined by equation (6). Let $A_t$ and $F_t$ denote the actual and forecast values, respectively, at data point t.

$$MAPE = \frac{1}{n} * \sum_{t=1}^{n} \left| (A_t - F_t)/A_t \right| \quad (6)$$

where:

n denotes the number of periods.

According to Na Na [13], the MAPE values can be classified as shown in Table 2.

**Table 2** MAPE value classification [13]

| Forecasting model performance | MAPE value |
|---|---|
| Very Good | < 10% |
| Good | 10% - 20% |
| Acceptable | 20% - 50% |
| Poor | > 50% |

In addition to MAPE, R-squared was computed to evaluate the goodness-of-fit between actual and forecasted values.

2) R-squared (The coefficient of Determination) is a number between 0 and 1 that measures the accuracy with which a model can anticipate a given result. Equation (7) shows how to calculate R-squared ($R^2$).

$$R^2 = 1 - SS_{regression}/SS_{total} \quad (7)$$

where:

$SS_{regression}$ is the regression sum of squares

$SS_{total}$ is the sum of all squares

### 3.6 Existing Studies on LSTM in Automotive Forecasting

LSTM networks have been applied to improve demand forecasting in automotive supply chains. Chandriah and Naraganahalli [14] reported enhanced accuracy in spare parts prediction using an LSTM model with a modified Adam optimizer, while Limbare and Agarwal [15] demonstrated similar benefits when integrating LSTM into demand planning and budgeting. Hybrid and advanced models further highlight LSTM's strengths. For instance, Suddala [16] found ARIMA–LSTM combinations outperform traditional ARIMA and Holt–Winters, and Terrada [17] emphasized LSTM's capacity to model nonlinear, volatile demand. In industrial contexts, Abbasimehr et al. [2] achieved consistent error reduction through optimized LSTM settings, and Cui et al. [18] introduced an attention-based bidirectional LSTM capable of capturing seasonality and trend effects. Na [13] also confirmed LSTM's superiority over classical forecasting techniques.

Despite these contributions, most applications focus on large-scale manufacturers, with limited attention to small or medium-sized vendors in emerging markets [17]. Comparisons with simple operational baselines such as the simple moving average (SMA) are also rare [15, 17]. Addressing these gaps, the present study evaluates LSTM against SMA using data from a local auto parts vendor in Samut Prakan, providing practical insights into its operational value.

## 4. Materials and methods

The proposed methodology is illustrated in Figure 2. The flowchart emphasises the main steps of our method.

### 4.1 Data Selection and Preprocessing

The sales order of different four parts are collected from an auto parts vendor in Samut Prakarn. These items were selected because they had the highest values. The data provided by the vendor were collected for three years from January 2023 to May 2025. The vendor also provides weekly sales reports for auto parts. Therefore, the data instance had 130 weeks, with 90 weeks for training and 40 weeks for testing. This dataset obtained no missing values or outliers.

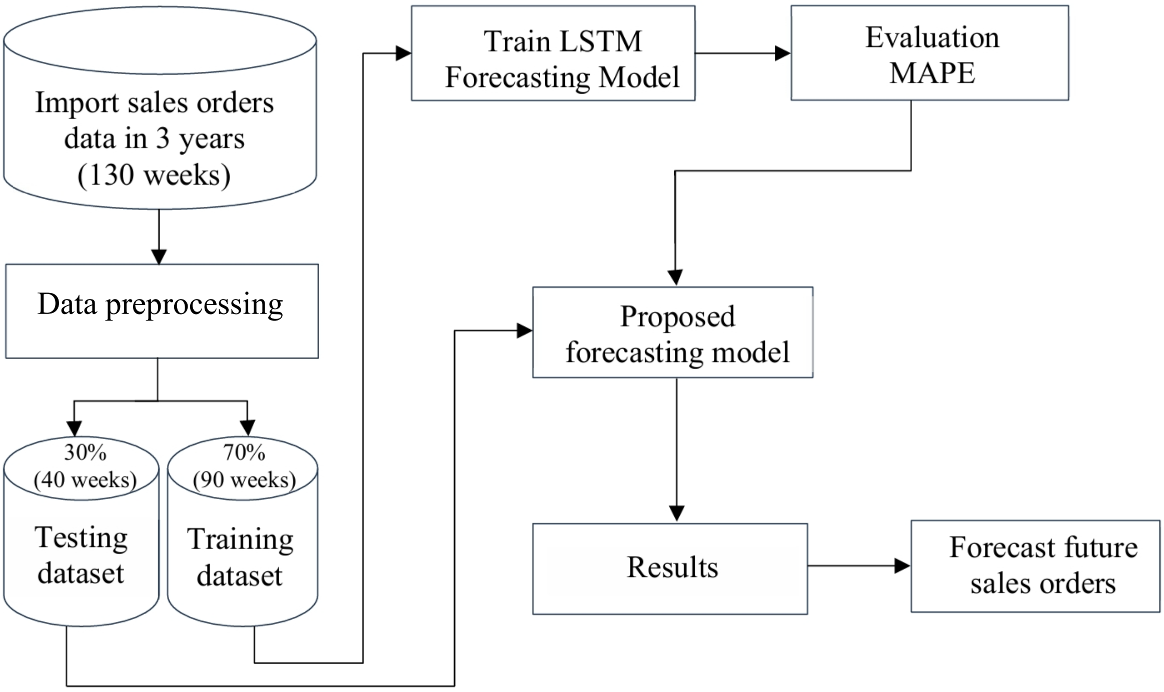


**Figure 2** Flowchart of the proposed forecasting method

Missing values were identified in Python. For time series with less than 5% missing data, linear interpolation was applied. Outliers were detected using the Interquartile Range (IQR) method, where values beyond Q1 – 1.5×IQR or Q3 + 1.5×IQR were flagged. These values were then visually inspected and replaced using median imputation if deemed erroneous.

**4.2 Goodness of fit test for LSTM network forecasting model**

The validation of the LSTM forecasting model was conducted by analysing the training and validation loss curves across four distinct auto parts. The evaluation focused on assessing the model's ability to generalize and avoid overfitting, as well as determining the optimal hyperparameters for the architecture. The initial model architecture was determined through a comprehensive grid search over a predefined validation set. The grid search evaluated various combinations of hyperparameters, including the number of hidden layers and units per layer. The optimizer and learning rate were also tuned to ensure optimal convergence. The results of the grid search indicated that the model with one hidden layer and 100 units per layer, using the Adam optimizer with a learning rate of 0.001, exhibited the fastest convergence and the lowest validation loss. However, the final model architecture was adjusted based on additional considerations to enhance performance and stability. The final configuration consisted of two hidden layers, each with 50 units, and utilized the ReLU activation function. This modification was made to balance the model's complexity and computational efficiency while maintaining its predictive accuracy. The goodness of fit for the LSTM model was evaluated by examining the convergence and stability of the training and validation loss curves, as illustrated in Figure 3. The curves demonstrated a consistent decrease in loss values during the initial epochs, indicating effective learning of the temporal dependencies in the sales data. The losses for all four auto parts converged to comparable values, suggesting that the model generalized well across different components.

Notably, the training and validation loss curves showed no signs of overfitting. Overfitting typically manifests as a divergence between training and validation performance, where the training loss continues to decrease while the validation loss plateaus or increases. In this case, both losses stabilized within a similar range, particularly after 10-15 epochs, indicating that the model effectively captured the underlying patterns without memorizing the training data. To further substantiate the model's performance, quantitative metrics were employed. The validation loss, measured using MAPE, was found to be consistently low across all four auto parts. Specifically, the MAPE values for the validation set were observed to be within a narrow range, demonstrating the model's robustness and reliability in forecasting sales for different components.

## 5. Results

This section compares the actual weekly sales against the predicted sales implemented by the LSTM model and the existing SMA-3weeks method as shown in Figure 4 below.

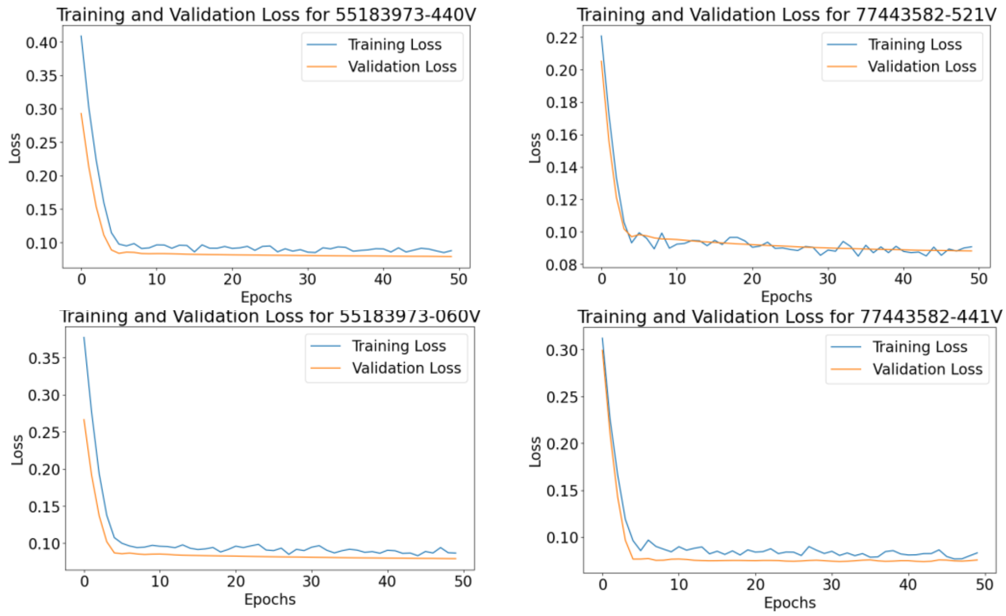


**Figure 3** Training and validation loss across four different auto parts

The blue line is the actual sales, which are very volatile with many peaks and trenches, reflecting irregular and dynamic sales data. The orange line represents the LSTM predictions, and the green line represents as the SMA-3weeks predictions. In Figure 4, the charts illustrate the predicted weekly sales for four auto parts, demonstrating that the LSTM model provided consistent predictions that followed the general sales trend. It generalises well across the dataset but is not overly sensitive to extreme fluctuations in actual sales. SMA-3weeks partially captured the trends, but delayed adapting to high increases or decreases in sales, resulting in larger deviations from actual values. The LSTM model predicted sales for sales order data and obtained the MAPE value.

**Table 3** The MAPE values from LSTM network

| Parts code | MAPE (%) | Performance |
|---|---|---|
| 77443582-441V | 8.75 | Very Good |
| 77443582-521V | 9.40 | Very Good |
| 55183973-060V | 10.07 | Good |
| 55183973-440V | 5.02 | Very Good |

To compare the models and their MAPE values, the accuracy of the proposed model and MAPE values were calculated. Table 3 shows that in all cases, the MAPE decreased.

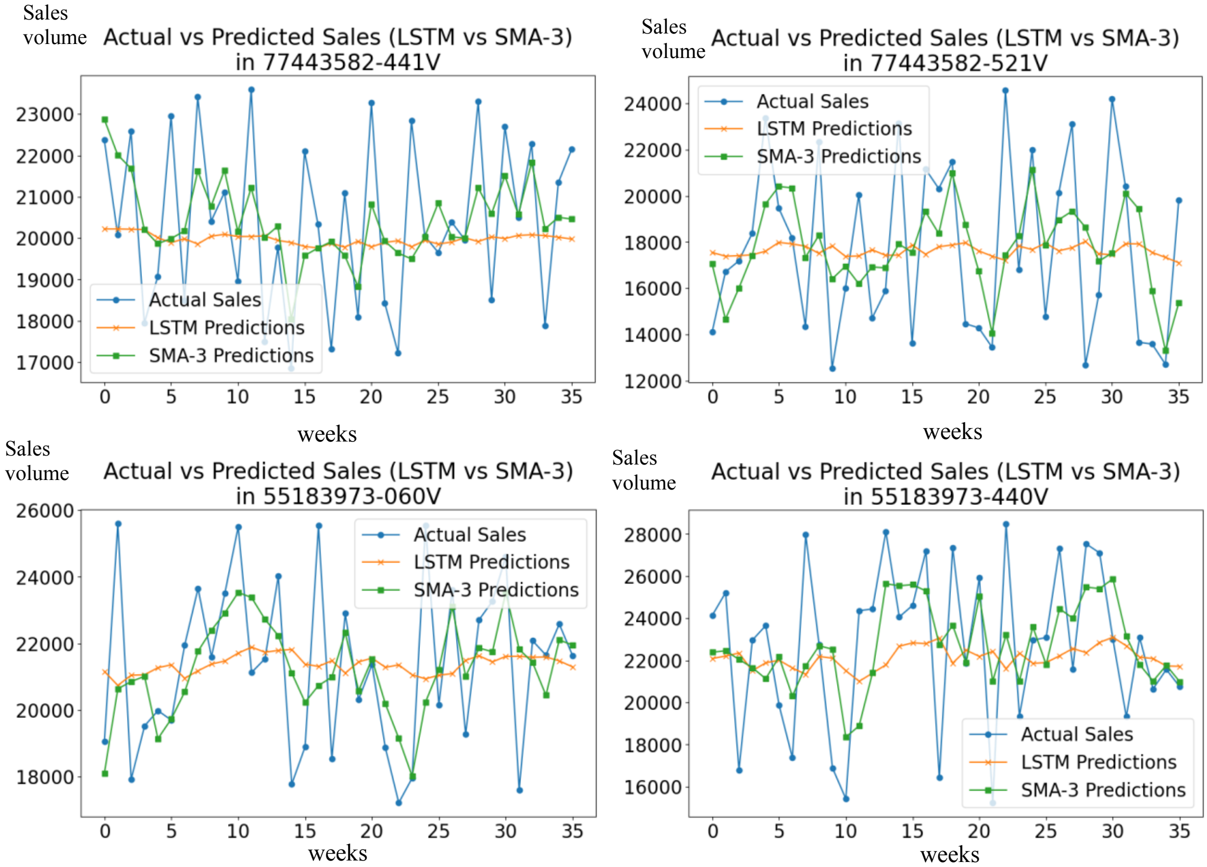


**Figure 4** Comparison of four auto parts: LSTM vs SMA-3weeks

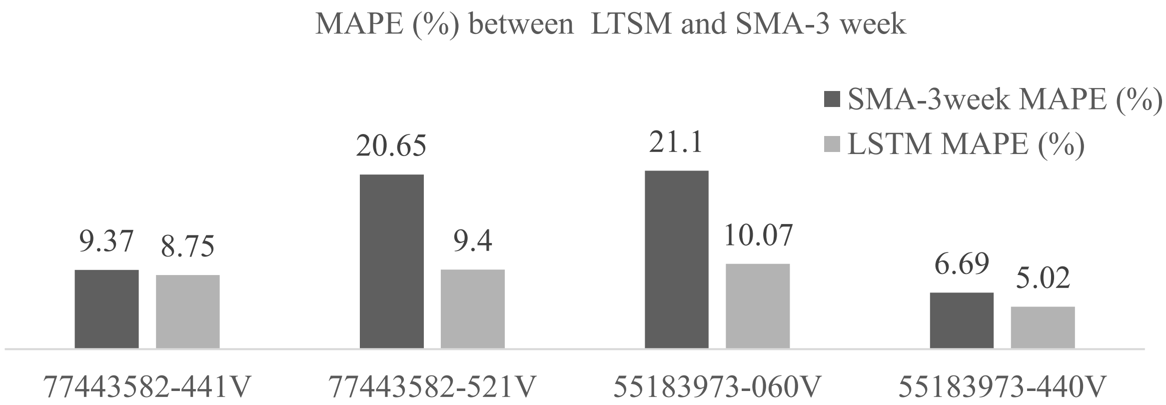


**Figure 5** Comparison of MAPE between LSTM and SMA-3week for four parts

Figure 5 shows the predictive errors of the LSTM model and existing 3-week Simple Moving Average (SMA). The average MAPE (%) improved by 6.14%. Compared to 77443582-441V, 77443582-521V, 55183973-060V, and 55183973-440V, the LSTM model reduced the MAPE by 0.62%, 11.25%, 11.03%, and 1.67%, respectively.

In addition, The LSTM model achieved the highest value where R-squared = 0.90228. It is concluded that the LSTM model achieved the highest accuracy and the lowest error rate.

## 6. Discussion and conclusion

The results show that the LSTM network is better for auto parts sales forecasting than the SMA-3weeks method. The training and validation losses indicate that the model is suitable for generalisation and convergence.

In addition, the MAPE evaluation showed an average improvement of 6.14%, and the model was good at capturing complex patterns and fluctuations in sales data. This agrees with the results reported by Pliszczuk et al. [19], LSTM models outperform traditional time-series methods by providing accurate predictions. These results prove that the LSTM network is a practical and efficient forecasting tool, especially for dealing with irregular and volatile demand in the auto parts industry. Based on the paper of Abbasimehr et al. [2], LSTM methods outperformed other time-series models when

## 7. Recommendation

In theoretical recommendations, a problem in sales forecasting occurs because the model over fits, especially when historical sales data are not sufficient to train. However, it also performs poorly on new data. Therefore, it is recommended to avoid this issue by applying fine-tuned hyper parameters, including early stopping, dropout regularisation, and monitoring validation loss.

In practical recommendations, it is recommended to improve the performance of LSTM models for sales forecasts, particularly for automobile parts vendors, through necessary improvements, as a solution to some of the extant limitations. This recommendation proposes the integration of data sources and historical sales data. Factors such as inventory levels, lead times, transport delays, and supplier performance provide insight into order sales patterns. As a future study, scholars will first aim to explore LSTM time series forecasting in other domains and regions to understand the differences in the sales forecasting mechanism. Second, only sales orders as structured data were considered, but unstructured data (emails, customer sentiment, and CRM activity logs) will also be deliberated to increase the accuracy of future sales forecasting. Finally, combining forecasting methods including linearity and non-linearity can be provided an accuracy rather than LSTM or SMA-3weeks will be constructed to test for forecasting model improvement.

## 8. References


[1] W. Liao, G. Y., Y. Yanchao, Y. Weijun, Y. Yaheng, and Z. Dongzhou, "Auto Parts Sales Prediction based on Machine Learning for Small Data and a Long Replacement Cycle, in *ACS/IEEE International Conference on Computer Systems and Applications*, November. 02-05, 2020, pp. 1-5.

[2] H. Abbasimehr, M. Shabani, and M. Yousefi, "An optimized model using LSTM network for demand forecasting," *Computers and Industrial Engineering*, vol. 143, pp. 1-13, May, 2020.

[3] J. Zheng, C. Xu, Z. Zhang, and X. Li, "Electric load forecasting in smart grids using Long-Short-Term- Memory based Recurrent Neural Network," in *51st Annual Conference on Information Sciences and Systems (CISS)*, Mar 22-24, 2017, pp. 1–6.

[4] C. Ghassen, H. Bakir, and H. Zaher, "E-commerce Time Series Forecasting using LSTM Neural Network and Support Vector Regression," in *Proceedings of the International Conference on Big Data and Internet of Things*, December. 20, 2017, pp. 80-84.

[5] B. Lim and S. Zohren, "Time-series forecasting with deep learning: a survey," *Philosophical Transactions of the Royal Society A: Mathematical, Physical and Engineering Sciences*, vol. 379, no. 2194, pp. 1-14, Feb, 2021.

[6] M. Khashei and M. Bijari, "A novel hybridization of artificial neural networks and ARIMA models

for time series forecasting," *Applied Soft Computing*, vol. 11, no. 2, pp. 2664-2675, Mar, 2011.

[7] K. Satashwa1, M. Nagre, and A. Bhange, "Artificial Neural network for effective demand forecasting of Product" in International Journal of Advances in *Engineering and Management (IJAEM)*, vol. 4, no. 10, pp. 982-986, Oct, 2022.

[8] A. R. S. Parmezan, V. M. Souza, and G. E. Batista, "Evaluation of statistical and machine learning models for time series prediction: Identifying the state-of-the-art and the best conditions for the use of each model," *Information Sciences*, vol. 484, pp. 302-337, Jan, 2019.

[9] H. A. Hurtado-Mora, A. H. García-Ruiz, R. Pichardo-Ramírez, L. J. González-del-Ángel, and L. A. Herrera-Barajas, "Sales Forecasting with LSTM, Custom Loss Function, and Hyperparameter Optimization: A Case Study," *Applied Sciences*, vol. 14, no. 21, p. 9957, Oct, 2024.

[10] A. El Filali, E. H. Lahmer, S. El Filali, M. Kasbouya, M. A. Ajouary, and S. Akantous, "Machine Learning Applications in Supply Chain Management: A Deep Learning Model Using an Optimized LSTM Network for Demand Forecasting," *International Journal of Intelligent Engineering and Systems*, vol. 15, no. 2, pp. 464-478, Feb, 2022.

[11] K. Greff, R. K. Srivastava, J. Koutník, B. R. Steunebrink, and J. Schmidhuber, "LSTM: A Search Space Odyssey," *IEEE Transactions on Neural Networks and Learning Systems*, vol. 28, no. 10, pp. 2222–2232, Mar, 2015.

[12] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Computation*, vol. 9, no. 8, pp. 1735-1780, Nov, 1997.

[13] N. Na, "Research on Supply Chain Demand Prediction Model Based on LSTM," *Procedia Computer Science*, vol. 243, no. C, pp. 313-322, Oct, 2024.

[14] K. Chandriah and R. Naraganahalli, "RNN/LSTM with modified Adam optimizer in deep learning approach for automobile spare parts demand forecasting," *Multimedia Tools and Applications*, vol. 80, no. 20, pp. 26145-26159, 2021.

[15] A. Limbare and R. Agarwal, "Demand forecasting and budget planning for automotive supply chain," *EAI Endorsed Transactions on Internet of Things*, vol. 10, no. 1, pp. 1-8, Nov. 2023.

[16] S. Suddala, "Dynamic demand forecasting in supply chains using hybrid ARIMA-LSTM architectures," *International Journal of Advanced Research*, vol. 12, no. 2, pp. 167-174, 2024.

[17] L. Terrada, M. Khaili, and H. Ouajji, "Demand forecasting model using deep learning methods for supply chain management 4.0," *International Journal of Advanced Computer Science and Applications*, vol. 13, no. 5, pp. 517-523, 2022.

[18] L. Cui, Y. Chen, J. Deng, and Z. Han, "A novel attLSTM framework combining the attention mechanism and bidirectional LSTM for demand forecasting," *Expert Systems with Applications*, vol. 254, Art.no. 124409, 2024.

[19] D. Pliszczuk, P. Lesiak, K. Zuk, and T. Cieplak, "Forecasting Sales in the Supply Chain Based on the LSTM Network: The Case of Furniture Industry," *European Research Studies Journal*, vol. 14, Special Issue 2, pp. 627-636, Jun, 2021.