Edu Komputika 12 (2) (2025)

Edu Komputika Journal

https://journal.unnes.ac.id/journals/edukom/

Comparison of Machine Learning and Deep Learning Algorithms for Daily Retail Sales Forecasting

Eko Purwanto[1)]✉, Bangun Prajadi Cipto Utomo[2)], Hanifah Permatasari[1)], and Farahwahida Mohd[3)]

[1]Department of Information System, Faculty of Computer Science, Universitas Duta Bangsa, Indonesia

[2]Department of Management, Faculty of Law and Business, Universitas Duta Bangsa, Indonesia

[3]Malaysian Institute Information Technology, Universiti Kuala Lumpur, Jalan Sultan Ismail 50250, Kuala Lumpur, Malaysia

Article Info

Article History:
Received: 22 August 2025
Revised: 8 January 2026
Accepted: 26 February 2026

Keywords:
Convolutional Neural Network, Deep Learning, Long Short-Term Memory, Machine Learning, Random Forest, Support Vector Machine.

Abstract

This study presents a comparative analysis of four machine learning (ML) and deep learning (DL) algorithms: Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM) for predicting daily retail sales time series. The models were evaluated using key metrics, such as Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared (R^2). Results show that RF and SVM outperformed both CNN and LSTM in terms of MAE (3500.28 and 3325.11, respectively) and RMSE (4660.60 and 4293.42, respectively). However, all models had negative R^2 values, indicating none could explain the variation in the data. LSTM, in particular, was the least efficient model, with an MAE of 54087.25, RMSE of 54257.51, and R^2 of -158.59. The poor performance of LSTM can be attributed to overfitting, improper model configuration, and misalignment with the nature of the data. The dataset used includes over 1,000 daily retail sales transaction records collected over one year, with key attributes like CustomerID, ProductID, Quantity, Price, TransactionDate, PaymentMethod, StoreLocation, ProductCategory, DiscountApplied, and TotalAmount. While the dataset is representative, its size and complexity may not have been sufficient for deep learning models like LSTM and CNN, which generally require larger datasets for optimal performance. This study highlights the challenges of using deep learning for retail forecasting and suggests future research should focus on refining models and incorporating external datasets to improve prediction accuracy.

✉ Correspondence Address:
Jl. Bhayangkara No. 55-57 Tipes, Serengan, Surakarta City, Central Java, 57155
E-mail: eko_purwanto@udb.ac.id

ISSN 2252-6811
E-ISSN 2599-297X

INTRODUCTION

Sales forecasting is one of the important decision-making tools in the retail industry, which influences inventory control, pricing policies, and demand forecasting. Having more accurate and reliable sales forecasts will help retailers to forecast market demands, ensure the right balance of stock, and ultimately enhance profitability. Traditional methods like moving averages and exponential smoothing are not always suitable for handling complex retail data, especially data with seasonal patterns and long-term relationships, as well as large volumes (Sharma et al., 2024). Such approaches fail to appropriately model the complicated structures in retail sales data, which highlights the need for advanced methodologies, such as Machine Learning (ML) and Deep Learning (DL), for increasing the forecasting accuracy (Ganguly & Mukherjee, 2024).

Recent works have indicated that ML methods, such as Random Forest (RF) and Support Vector Machines (SVM), are capable of capturing non-linear relationships in high-dimensional feature spaces, and can enable a variety of predictive tasks in the retail (Ganguly & Mukherjee, 2024). But these models usually cannot capture the temporal relationships and the sequential features in time series prediction. On the other hand, DL models such as Convolutional Neural Networks (CNNs) and Long Short-Term Memory (LSTMs) have shown better performance in learning sequential patterns and processing massive datasets (de Castro Moraes et al., 2024; Fukai et al., 2024; Khine, 2023). While there has been progress on implementing such algorithms in retail sales forecasts, there are few studies that compare how well such machine learning and deep learning models work in retail sales prediction in such point-of-sale environments. It is necessary to compare these two algorithms so as to analyze which one is more capable of learning the intricate structures of retail data.

Investigated that Random Forest and Support Vector Machines function better compared to CNN, LSTM for retail sales Prediction, especially when no equivalent feature is used. Nonetheless, showed that LSTM works better with complex and big datasets. These results emphasize the importance of ongoing research into systematic comparison of ML and DL models (Nasseri et al., 2023; Shak et al., 2024).

However, although ML and DL for retail sales forecasting have gained momentum, little research has been done on comprehensive comparisons between algorithms, especially in terms of predictive accuracy, computational efficiency, and robustness (Ahmed et al., 2024; Pustokhina & Pustokhin, 2023). There are also a few reports comparing different ML and DL methods for daily retail sales forecasting directly, and generally are either special comparisons of different methods or comparisons of limited data categories (Eglite & Birzniece, 2022; Haque et al., 2023). This gap is what the present study has tried to bridge by evaluating four algorithms, namely RF, SVM (ML) and CNN, LSTM (DL) applied on a daily retail sales data using multiple evaluation metrics like Mean Absolute Error (MAE), Root Mean Squared Error (RMSE) (Naskinova et al., 2024).

The research purpose is to compare and analyze the performances of machine learning and deep learning models – RF, SVM, CNN, and LSTM in the daily retail sales forecasting, and determine which of these models can be a better model to predict the daily retail sales precisely and most effectively. Understanding the pros and cons of such algorithms to work on retail data is important for generating insights that are actionable by retail professionals.

In the domain of daily retail sales forecasting, we will systematically compare four different algorithms, including RF, SVM (machine learning) and CNN, LSTM (deep learning). The performance of the algorithms will also be evaluated in terms of their forecasting performance, computational time, and robustness, using well-known measures of accuracy, such as MAE and RMSE (Fukai et al., 2024). This extensive comparison will also shed light on which model would be better for the retail business, resulting in a sound decision for practitioners to choose the forecasting model (de Castro Moraes et al., 2024).

The projected impact of this research is the discovery of the best algorithms for day-to-day retail sales prediction in terms of both predictive performance, efficiency, and robustness. This project will enrich the theoretical understanding of time series forecasting techniques and also bring the practical implications to the retail industry through the comparison of machine learning and deep learning models. Better sales prediction directly translates into better decision making, better inventory management, unit-level profitability, and better operating efficiency - something that is a "killer app" for retailers.

RESEARCH METHODS

A. Dataset

In this study, the authors use the daily retail sales information with transaction details of customers, products, and transactions for our analysis. The dataset --from retail transaction-- covers one year and comprises 1000+ lines of data. The key attributes in this dataset are CustomerID, ProductID, Quantity, Price, TransactionDate, PaymentMethod, StoreLocation, ProductCategory, DiscountApplied, and TotalAmount. These attributes make up a holistic view of day-to-day retail transactions, capturing the number and amount of sales and factors such as promotions, payment types, and store locations.

B. Data Preprocessing

Prior to training the models, several preprocessing operations are applied to polish the data according to the needs of machine learning. At first, the missing values in the numerical features are imputed through a mean imputation. This is then followed by the decomposition of the TransactionDate into Date and Time components to represent the temporal dynamics of the transactions. In addition, daily sales are calculated by summing up the Quantity and TotalAmount at the day level, allowing the models to focus on general sales patterns. Numerical attributes such as Price, Quantity, and Total Amount are MinMaxScaled to bring them to the same scale. Binary encoding is utilized to encode categorical features such as PaymentMethod and ProductCategory. The data is finally split into 80% for training purposes and 20% for testing purposes by ensuring the testing set contains never-seen-future data for evaluation (Chebli et al., 2024; Pindiyan & M., 2024).

C. Modeling

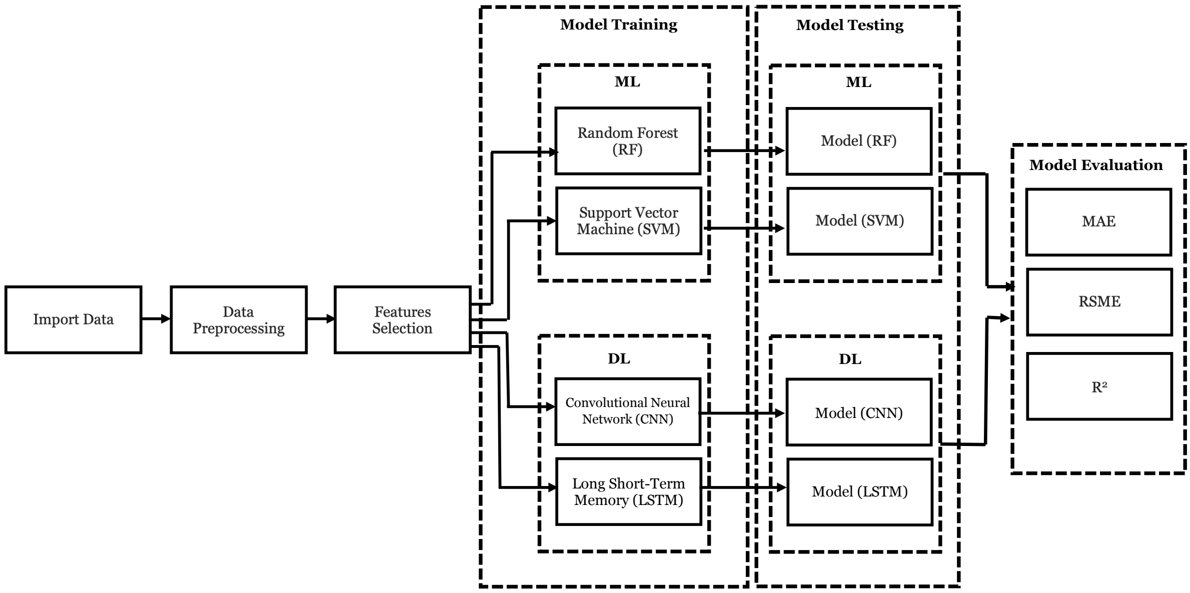

Figure 1. Proposed model

In this study, we compare the performance of four algorithms for daily retail sales time series forecast, namely, RF, SVM, CNN, and LSTM. All the methods were selected as they have their own strengths and were implemented in the analysis for particular dataset characteristics (mega-size, non-linearity, temporal relations, feature representations).

1. Random Forest

Random Forest (RF) is an ensemble learning method that uses multiple decision trees and random selections of the data to generate them. It combines the predictions of all the trees to give a final output, making it a robust method for working with complex data (J. Zhao et al., 2024). The prediction function of Random Forest is expressed by Equation 1.

$$f(x) = \frac{1}{N}\sum_{i=1}^{N} h_i(x) \tag{1}$$

where $h_i(x)$ represents the prediction from the i-th tree, and N is the total number of trees in the forest. By averaging the outputs of individual trees, Random Forest reduces the risk of overfitting, providing more stable predictions.

2. Support Vector Machine

The Support Vector Machine (SVM) is a strong model used in classification. It is also quite effective in regression as well, once it is generalized to SVR. Support Vector Machine (SVM) method searches for an optimal hyperplane that maximizes the margin to separate instances of each class in a multi-dimensional space. In SVR, we aim to reduce the prediction error while maintaining a large margin (Suzuki, 2020). The prediction function of SVR is shown in Equation 2.

$$f(x) = \langle w, x\rangle + b \tag{2}$$

where $\langle w, x\rangle$ is the dot product of the weight vector w and input vector x, and b is the bias. The model's objective is to minimize the squared norm of the weight vector, subject to the constraint that the prediction error is within a specified tolerance:

$$min\ \frac{1}{2}\ \|w\|^2\ subject\ |y_i - f(x_i)| \leq \epsilon + \xi_i \quad (3)$$

where y_i is the actual value, $f(x_i)$ is the predicted value, and ϵ is the margin of tolerance for errors.

3. Convolutional Neural Network

A Convolutional Neural Network (CNN) is good at handling high-dimensional input and extracting complicated patterns using convolutional layers and pooling layers. (i) In CNNs, the features are extracted using a convolution operation, and then pooling is performed to reduce their dimension (Cao & Sun, 2021). The convolution operation is expressed in Equation 4.

$$S(i,j) = (I * K)(i,j) = \sum_m \sum_n I(i+m, j+n)K(m,n) \quad (4)$$

where $S(i,j)$ is the result of the convolution operation at position (i,j), and $K(m,n)$ is the kernel or filter. The max pooling operation, which reduces the spatial dimension of the feature map, is defined as Equation 5.

$$P(i,j) = \max_{m,n \in POOLING\ WINDOW} I(i+m, j+n) \quad (5)$$

This allows CNNs to detect complex temporal patterns and hierarchical features within the dataset, which are crucial for accurately predicting sales trends.

4. Long Short-Term Memory

Long Short-Term Memory (LSTM) is a special type of Recurrent Neural Network (RNN) capable of learning long-range dependencies in temporal input. LSTM is very good at time series prediction, due to its ability to preserve information across long periods and thus avoid the vanishing gradient problem in traditional RNNs (Ravikumar & Sriraman, 2023). The information propagating through the LSTMs is regulated by gates, and the primary operations are Equations 6 to 10.

Input Gate:

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (6)$$

Forget Gate:

$$f_t = \sigma\left(W_f \cdot [h_{t-1}, x_t] + b_f\right) \quad (7)$$

Output Gate:

$$o_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (8)$$

Cell State Update:

$$c_t = f_t \cdot c_{t-1} + i_t \cdot tanh(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (9)$$

Hidden State Update:

$$h_t = o_t \cdot tanh(c_t) \quad (10)$$

These gates enable the LSTM to capture both short- and long-term dependencies in the input, which is crucial for identifying intricate patterns in time series data.

D. Model Training and Evaluation

All architectures are trained with the Adam Optimizer, the learning rate of which is adjusted during training. Grid Search is utilized to optimize the hyperparameters of each model (including C and epsilon for SVM and the number of layers and the number of units for LSTM) (Makinde, 2024). The evaluations are carried out with the metrics in Equations 11 through 13.

Mean Absolute Error (MAE):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - y_i^{\wedge}\right| \quad (11)$$

Where y_i is the actual value and $y_i^{\wedge}$ is the predicted value.

Root Mean Squared Error (RMSE):

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i - y_i^{\wedge})^2} \quad (12)$$

R-Squared (R^2)

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - y_i^{\wedge}\right)^2}{\sum_{i=1}^{n}\left(y_i - y_i^{-}\right)^2} \quad (13)$$

Where y_i^{-} is the mean of the observed data.

We also include the training time and the evaluation time for each model to evaluate computation efficiency. The MAE, RMSE, and training time of models will be compared in terms of accuracy and research time after training and testing. The model that has the lowest MAE and RMSE and the quickest training time for forecasting daily retail sales will be treated as the most effective model.

E. Model Parameter

To provide a clearer understanding of the model configurations used in this study, Table 1 below summarizes the key parameters for each of the models applied (Random Forest, SVM, CNN, and LSTM), including those used in the hybrid ensemble approach.

Table 1. Model Parameter Summary

Model	Key Parameters	Description
RF	n_estimators=50, random_state=42	An ensemble learning method using decision trees for more stable predictions.
SVM	C=20, epsilon=0.2	A regression model that finds the optimal hyperplane to minimize prediction errors.
CNN	Filter=16, Kernel Size=2, Activation='relu', Epochs=5, Batch Size=8	Used for processing sequential data like time series with convolutional layers.
LSTM	Units=24, Epochs=5, Batch Size=8	A recurrent neural network model designed to handle long-range sequential data.

RESULT AND DISCUSSION

Performance of the four models: Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Long Short- Term Memory (LSTM) models was evaluated using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and R-squared (R^2). A performance comparison of the different models is summarized in Table 2. The Random Forest model had good performance with an MAE of 3500.28 and an RMSE of 4660.60. The model wasn't even close, with a negative R^2 score of -0.18, suggesting that even though the model explained some amount of error, it failed to properly capture any underlying patterns in the sales data. This is consistent with previous findings [20], which indicated that the ensemble models, such as Random Forest, may be stable in the prediction but may lack the flexibility of fully characterizing the time-series information.

Table 2. Comparative Analysis of Machine Learning and Deep Learning Algorithms

Model	MAE	RMSE	R^2
Random Forest	3500.279764	4660.604339	-0.177504
SVM	3325.106250	4293.419488	0.000726
CNN	3534.713576	4717.297588	-0.206325
LSTM	54087.252078	54257.512688	-158.587049

The performance of the Support Vector Machine (SVM) was a little better with a MAE of 3325.11 and RMSE of 4293.42. Nevertheless, the R^2 value of 0.0007 indicates that the SVM only poorly reflected the trend of the data, so its ability for prediction was unsuccessful. These observations are in agreement with previous work (Kaneko, 2021; Velasco et al., 2020), which reported that though

with works, such as (Murray et al., 2022; Wibawa et al., 2022), where, in general, CNNs (which perform very well at image-based learning) failed at this task or, also, when not properly adapted to time series forecasting.

Even though it was created to work with sequential data, LSTM had the lowest result, with a MAE of 54087.25 and RMSE of 54257.51. The R^2 of -158.59 indicates that the model failed to model the temporal relationships on the sales. This bad performance can be explained by overfitting, wrong model configuration, and a model architecture that does not fit with the nature of the data. This result differs from the research in (de Castro Moraes et al., 2024; Khine, 2023; Y. Zhao et al., 2021), since the research in those papers demonstrated very high forecasting performance of LSTM, which means that the model may not be a good model for retail sales forecasting without strong tuning.

, including using larger datasets or incorporating relevant external features, such as promotions or economic conditions, that may influence sales.

SVM minimizes errors well, it is less powerful in expressing the complex temporal dependencies required by retail sales forecasting.

On the other hand, CNN was inferior to it, with a MAE of 3534.71 and RMSE of 4717.30. The negative R^2 score of -0.21 indicates that CNN is unable to capture time dynamics in the data. This is consistent

Although the CNN and LSTM models are designed to handle time-series data, both showed lower performance compared to RF and SVM. One main reason is the limitations of the dataset, which was not large or diverse enough to fully leverage the capabilities of deep learning models, especially LSTM, which requires data with clearer long-term dependencies. Additionally, overfitting became an issue for the LSTM model, as the limited number of epochs (5 epochs) caused the model to overfit the training data and fail to generalize to the test data. CNN also struggled to capture complex time-series patterns, as its architecture is better suited for grid-like data, such as images, and the filters used may have been too small to capture the daily sales patterns. To improve the performance of both models, further optimization is needed

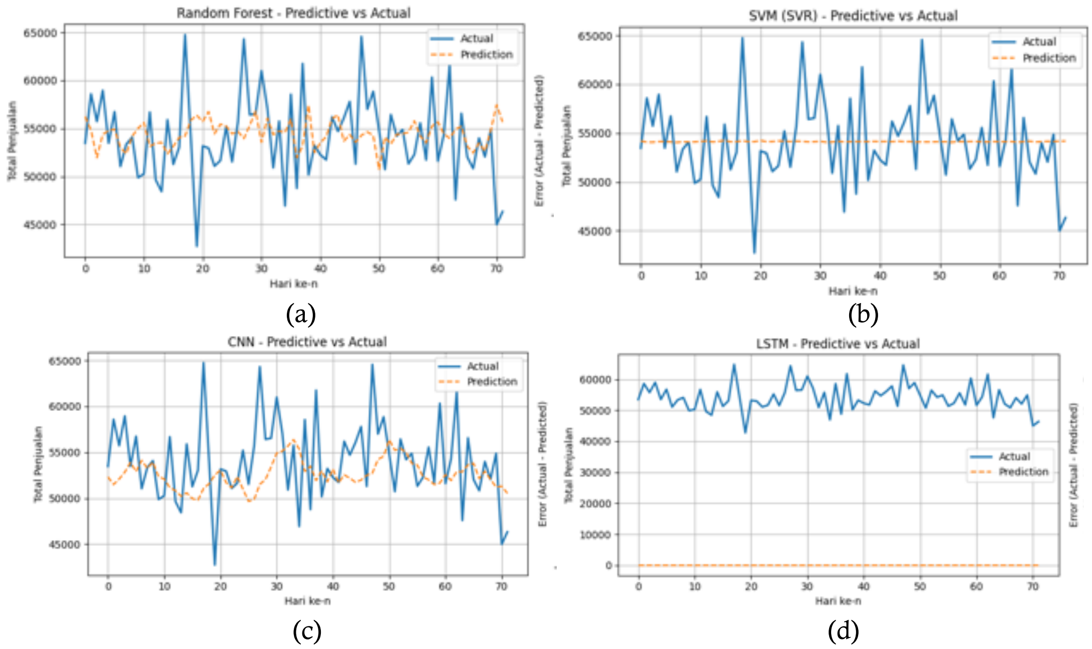

Figure 2. Comparison models predictive vs actual plot (a) random forest (b) support vector machine (c) convolutional neural network (d) long short-term memory

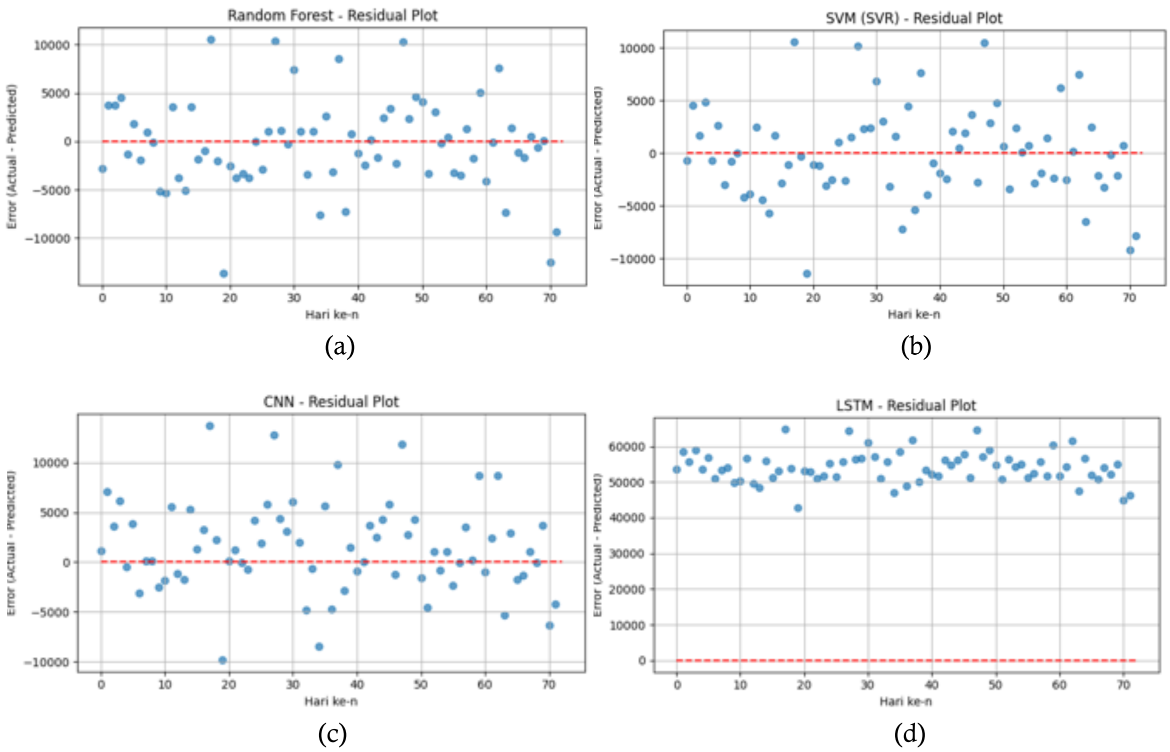

Figure 3. Comparison models residual plot (a) random forest (b) support vector machine (c) convolutional neural network (d) long short-term memory

The Prediction vs Actual and Residual Plots (Figures 2 and 3) also highlight the performance of the models. Random Forest and SVM were good at predicting, and observed sales trends increased the accuracy of prediction, while the predictive accuracy was reduced during high-sales periods. CNN, on the other hand, had very scattered predictions, which suggests its inability to model the data properly, and LSTM had a very big gap between true and predicted values, which also reveals the same. From the Residual Plots, the residual points of Random Forest and SVM were more evenly distributed, which indicated that the Random Forest model and the SVM model have better captured the trends of sales. CNN and LSTM, on the other hand, showed high spikes and noisy residuals, as can be seen in figures 1 (a), (b), (c), and (d), which means that those models did not sufficiently fit the data.

The findings of this research reveal that Random Forest and SVM were the best performing models predicting daily sales per store and per department, namely, since the R^2 values were negative, the models did not explain the variation, making the predictions with the values of the attributes in the training set. This indicates that although these models can reduce error, they are not adequate to understand the intricate patterns in retail sales. This result is consistent with existing research that states classical machine learning models, such as Random Forest and SVM, are preserving adequate performance on retail forecasting, and their limitation is in decoding data variability (Ganguly & Mukherjee, 2024; Pustokhina & Pustokhin, 2023; Sharma et al., 2024; Yu, 2024).

Poor performance of CNN and LSTM causes concerns about the adaptability of deep learning models for sales prediction in retail. While deep learning models such as LSTM are designed to handle long-range dependencies in sequences, the results from this study showed that LSTM was not able to capture the needed temporal relationships, possibly due to overfitting, lack of data, or an incorrectly calibrated model. Also, CNN also failed to give a proper prediction which demonstrates that more meticulous tuning will be required for these models in time series jobs. These findings validate (Fukai et al., 2024; Meenakshi, 2024; Shak et al., 2024), in which deep learning models were observed to yield limited performance for time series forecasting with no significant optimization and preprocessing.

The findings in this study have significant implications for the retail sector. Random Forest and SVM can also be used to take accurate sales predictions and help in inventory management, demand forecasting and promotional planning. However, the moderate degree to which these models can account for data variance means that integrating more sources of data (eg, economic models, competitor pricing, and customer sentiment) could improve the quality of data, resulting in improved forecast accuracy.

In the future, we could work on optimizing the model performance by hyperparameter tuning, trying other advanced models such as XGBoost and Prophet, and adding more external sources to increase prediction precision. Second, LSTM and CNN may also be re-tested with different structures or deeper hyperparameter search, to see the possibility of whether they can be adjusted to retail sales forecasting tasks.

CONCLUSION

In this paper, a comprehensive comparative study is conducted on four commonly used machine learning and deep learning algorithms, namely, Random Forest (RF), Support Vector Machine (SVM), Convolutional Neural Network (CNN), and Long Short-Term Memory (LSTM), for predicting daily retail sales with a time series. The results show that the Random Forest and SVM outperformed the CNN and LSTM for MAE and RMSE. Although the model efficiency was good in terms of error reduction, the negative R^2 values of all models indicated that none of the algorithms considered fully absorbed the explained variance of their underlying retail sales data, implying only a relative predictive capability of the models. In general, Random Forest performed the best, followed closely by SVM.

CNN and LSTM models, despite their capacity to handle sequential data, disagreed largely in predictions, particularly in peak demand periods. This demonstrates the difficulty that deep learning models could incur while trying to forecast time series data in retail due to the wide variation in patterns in the non-linear fashion in retail data, as deep learning models might fail to capture characteristics of the non-linear patterns without customization.

This study has several practical implications for the retail environment. Random Forest and SVM are reliable to predict daily retail sales, but deep learning models (CNN and LSTM) still need further tuning to excel in this domain. Retailers may find value in employing Random Forest and SVM in sales prediction tasks; however, the negative R^2 values suggest the necessity of model improvement, potentially by better feature engineering, hyperparameter tuning, and leveraging external data (economic indicators, consumer behavior).

In future work, the task would be to attempt to enhance the power of deep learning models through various architectures, training strategies, and hyperparameter settings. Also, the inclusion of external data (e.g., weather, promotion, economy) would further improve the forecasting performance of the model. Advanced machine-learning techniques, such as XGBoost or Prophet, may be of interest in improving model predictions as well.

ACKNOWLEDGEMENTS

This investigation was only made possible due to the crucial support from several entities. We would like to thank the Ministry of Higher Education, Science, and Technology (Kemendiktisaintek) for the financial support, substantially provided for the successful completion of the study. This research could not have been done without their help. The authors would also like to thank Universitas Duta Bangsa for providing academic resources, facilities, and support infrastructure for this research. The university's institutional support was essential for the success of this project. We would also like to express our gratitude to SG Komputer for being a strategic partner, who has provided us with data and equipment that are vital for the study of SMEs. Their help was crucial for securing the methodological quality and empirical soundness of the research. Lastly, the authors would like to thank their families for their tremendous support and understanding in the course of this investigation.

REFERENCES

Ahmed, R., Hasnain, M., Mahmood, M. H., & Mehmood, M. (2024). *Comparison of Deep Learning Algorithms for Retail Sales Forecasting*. *1*(3), 112–126. https://doi.org/10.62762/tis.2024.300700

Cao, X., & Sun, J. (2021). *Convolutional Neural Networks (CNN).* 83–109. https://doi.org/10.1007/978-3-030-82184-5_6

Chebli, A., Daas, S., & Hafs, T. (2024). Evaluating the impact of data imputation on model precision in machine learning. *Studies in Engineering and Exact Sciences*, *5*(2), e8310. https://doi.org/10.54021/seesv5n2-256

de Castro Moraes, T., Yuan, X.-M., & Chew, E. P. (2024). Hybrid convolutional long short-term memory models for sales forecasting in retail. *Journal of Forecasting*. https://doi.org/10.1002/for.3073

Eglite, L., & Birzniece, I. (2022). Retail Sales Forecasting Using Deep Learning: Systematic Literature Review. *Complex Systems Informatics and Modeling Quarterly*, *30*(30), 53–62. https://doi.org/10.7250/csimq.2022-30.03

Fukai, F. M., Cavalieri, D. C., & de Castro, F. Z. (2024). *Long short-term memory neural networks applied in demand forecast in the retail market*. https://doi.org/10.55592/cilamce.v6i06.10158

Ganguly, P., & Mukherjee, I. (2024). *Enhancing Retail Sales Forecasting with Optimized Machine Learning Models*. https://doi.org/10.48550/arxiv.2410.13773

Haque, M. S., Amin, M. S. A., & Miah, J. (2023). *Retail Demand Forecasting: A Comparative Study for Multivariate Time Series*. https://doi.org/10.21203/rs.3.rs-3280263/v1

Kaneko, H. (2021). Support vector regression that takes into consideration the importance of explanatory variables. *Journal of Chemometrics*, *35*(4). https://doi.org/10.1002/CEM.3327

Khine, M. T. (2023). Retail Demand Forecasting Using Sequence to Sequence Long Short-Term Memory Networks. *International Conferences on Computing Advancements*, 208–213. https://doi.org/10.1109/ICCA51723.2023.10181450

Makinde, A. S. (2024). *Optimizing Time Series Forecasting: A Comparative Study of Adam and Nesterov Accelerated Gradient on LSTM and GRU networks Using Stock Market data*. https://doi.org/10.48550/arxiv.2410.01843

Meenakshi, M. (2024). Enhanced Retail Sales Forecasting Using Optimized CNN-LSTM Hybrid Model: A Time Series-Based Analysis. *Deleted Journal*, *32*(2s), 514–537. https://doi.org/10.52783/cana.v32.2485

Murray, C., Chaurasia, P., Hollywood, L., & Coyle, D. (2022). *A Comparative Analysis of State-of-the-Art Time Series Forecasting Algorithms*. 89–95. https://doi.org/10.1109/csci58124.2022.00021

Naskinova, I., Kolev, M., & Lazarova, M. (2024). Forecasting Strategies in Retail: Utilizing Advanced Machine Learning Methods while Safeguarding Privacy. *Journal of Physics*, *2910*(1), 12008. https://doi.org/10.1088/1742-6596/2910/1/012008

Nasseri, M., Falatouri, T., Brandtner, P., & Darbanian, F. (2023). Applying Machine Learning in Retail Demand Prediction—A Comparison of Tree-Based Ensembles and Long Short-Term Memory-Based Deep Learning. *Applied Sciences*. https://doi.org/10.3390/app131911112

Pindiyan, A. V., & M., P. (2024). *Machine Learning-Based Imputation Techniques Analysis and Study*. 1–6. https://doi.org/10.1109/iceect61758.2024.10739330

Pustokhina, I. V, & Pustokhin, D. A. (2023). A Comparative Analysis of Traditional Forecasting Methods and Machine Learning Techniques for Sales Prediction in E-commerce. *American Journal of Business and Operations Research*, *10*(2), 39–51. https://doi.org/10.54216/ajbor.100205

Ravikumar, A., & Sriraman, H. (2023). *A Deep Understanding of Long Short-Term Memory for Solving Vanishing Error Problem* (pp. 74–90). IGI Global. https://doi.org/10.4018/978-1-6684-8531-6.ch004

Shak, M. S., Mozumder, M. S. A., Hasan, M. A., Das, A. C., Miah, M. R., Akter, S., & Hossain, M. N. (2024). *Optimizing retail demand forecasting: a performance evaluation of machine learning models including lstm and gradient boosting*. https://doi.org/10.37547/tajet/volume06issue09-09

Sharma, H., Kumar, D., Baskar, S., Shukla, D., Singh, D., Tiwari, S., & Danish, M. (2024). Enhancement of Sales Forecasting and Prediction with Machine Learning Methods. *International Research Journal of Computer Science*, *11*(11), 641–644. https://doi.org/10.26562/irjcs.2024.v1111.02

Suzuki, J. (2020). *Support Vector Machine* (pp. 171–192). Springer, Singapore. https://doi.org/10.1007/978-981-15-7568-6_9

Velasco, L. C. P., Polestico, D. L. L., Abella, D. M. M., Alegata, G. T., & Luna, G. C. (2020). *Performance Analysis of Support Vector Regression Machine Models in Day-Ahead Load Forecasting* (pp. 541–555). Springer, Cham. https://doi.org/10.1007/978-3-030-39442-4_40

Wibawa, A. P., Utama, A. B. P., Elmunsyah, H., Pujianto, U., Dwiyanto, F. A., & Hernandez, L. (2022). Time-series analysis with smoothed Convolutional Neural Network. *Journal of Big Data*, *9*(1). https://doi.org/10.1186/s40537-022-00599-y

Yu, C. (2024). Walmart Sales Forecasting using Different Models. *Highlights in Science Engineering and Technology*, *92*, 302–307. https://doi.org/10.54097/kqf76062

Zhao, J., Lee, C.-D., Chen, G., & Zhang, J. (2024). *Research on the Prediction Application of Multiple Classification Datasets Based on Random Forest Model*. 156–161. https://doi.org/10.1109/icpics62053.2024.10795875

Zhao, Y., Ren, X., & Zhang, X. (2021). *Optimization of a Comprehensive Sequence Forecasting Framework Based on DAE-LSTM Algorithm*. *1746*(1), 12087. https://doi.org/10.1088/1742-6596/1746/1/012087