INTERNATIONAL JOURNAL OF ADVANCED COMPUTING AND INFORMATICS
VOLUME 2, ISSUE 2, 2026, pp. 75 – 82
ISSN: 3089-7483, DOI: https://doi.org/10.71129/ijaci.v2i2.pp75-82

RESEARCH ARTICLE

Hybrid LSTM-MLP Framework for Enhanced Sales Forecasting

Onesinus Saut Parulian Tamba [a]*, Md. Munirul Hasan [b]

[a]Faculty of Information Technology, Universitas Nusa Mandiri, 12540 Jakarta, Indonesia
[b]Faculty of Computing, University Malaysia Pahang Al-Sultan Abdullah, Pekan 26600, Malaysia

ABSTRACT - This study introduces a hybrid framework combining Long Short-Term Memory (LSTM) networks and Multi-Layer Perceptrons (MLP) to enhance sales forecasting through advanced feature processing. The LSTM component is employed to capture temporal dependencies in sequential sales data, while the MLP focuses on processing numerical and categorical features to complement the predictive accuracy. This integration leverages the strengths of both architectures, enabling more comprehensive and effective feature processing compared to traditional approaches. The methodology encompasses data preprocessing, model development, and evaluation across performance metrics, including Mean Absolute Error (MAE), Mean Squared Error (MSE), Root Mean Squared Error (RMSE), and the coefficient of determination (R2). By employing real-world sales datasets, the framework undergoes rigorous testing to compare its effectiveness against existing models. Experimental results show that the hybrid LSTM-MLP model achieves notable improvements, including reduced MAE, MSE, and RMSE, while delivering 99% R2 values, higher than standalone LSTM, MLP, and other traditional methods. These findings underscore the model's potential to offer accurate and stable predictions, demonstrating its utility for forecasting tasks where tabular and sequential data interplay. The proposed framework not only contributes to advancing predictive methodologies in sales forecasting but also serves as a reference for adopting hybrid strategies in similar studies. Its effectiveness in handling feature complexity and producing reliable forecasts highlights its applicability in real-world forecasting scenarios and lays the foundation for further research in hybrid neural network designs.

Keywords: Sales Forecasting; LSTM; MLP; Hybrid Model; Feature Processing

1. Introduction

Accurate sales forecasting is pivotal in driving business success, affecting areas such as inventory management, supply chain optimization, and strategic planning [1]. Reliable forecasts enable companies to reduce operational costs, minimize stockouts or overstock situations, and align production schedules with market demands. However, the complexity of modern commerce introduces challenges in obtaining accurate predictions due to fluctuating market conditions, seasonality, and the intricate interplay of various external factors [2].

Traditional statistical methods, such as ARIMA or simple moving averages, effectively model linear trends but often fail to capture non-linear relationships and temporal dependencies in real-world sales data [3][4]. Similarly, standalone machine learning models like Gradient Boosting Machines (GBM) or basic neural networks struggle to simultaneously process sequential patterns and rich tabular data, leading to suboptimal outcomes [5]. The advent of deep learning techniques like Long Short-Term Memory (LSTM) networks and Multi-Layer Perceptrons (MLP) has created opportunities to overcome these limitations. LSTM is specifically designed for capturing long-term dependencies in time-series data, while MLP excels in analyzing tabular inputs, such as categorical and numerical features [6].

While these architectures offer considerable potential, relying solely on a single method limits the ability to capture complex interdependencies present in both temporal and static features. Recent advancements in hybrid approaches show promise by integrating complementary strengths of various models, but their application in sales forecasting remains underexplored [7]. Hybrid frameworks, particularly combining LSTM and MLP, stand out as a viable solution for addressing the gaps between temporal and feature-based prediction techniques [8].

This study focuses on developing a robust hybrid model that leverages the strengths of LSTM for sequential data and MLP for structured features. The integration aims to provide a comprehensive solution, yielding accurate and efficient sales forecasts. By bridging existing limitations in single-method approaches, this framework is expected to contribute significantly to sales forecasting accuracy while reducing computational overhead, offering practical utility across industries like retail, e-commerce, and manufacturing.

Received 25 February 2025, Accepted 5 September 2025
Available online 14 October 2025

* **Corresponding Author**
E-mail: onesinus231@gmail.com (Onesinus Saut Parulian Tamba)
© The Author(s) 2026. This is an open-access article under the CC BY license (https://creativecommons.org/licenses/by/4.0/)

2. Related Works

The challenge of optimizing Cross-Border E-commerce (CBEC) supply chain management, particularly in producing accurate sales forecasts and allocating resources in a dynamic and complex environment, led to the proposal of a novel AI-based framework designed to enhance efficiency and profitability [9]. This proactive system overcomes the constraints of traditional reactive approaches by utilizing a unique combination of methods: Neural Network-based Demand Forecasting is performed by a custom neural network model, whose weight vector is adjusted using a black-hole optimization (BHO) algorithm; subsequently, a Fuzzy Logic model assesses resource renewal needs by combining the predicted demand values with current inventory and sales features; and finally, a sealed-bid auction model with a first-price mechanism is employed to evaluate the value of each seller for resource allocation, ultimately determining the optimal distribution pattern for limited supplier resources. The effectiveness of this methodology was demonstrated by superior performance results compared to other models (SARIMA, CNN, ARIMA-NARNN), revealing reduced prediction errors with RMSE and MAE values of 22.31 and 18.76, respectively, and achieving the highest accuracy of 94.8911% in predicting average sales values. This superior performance is largely attributed to the innovative combination of fuzzy logic for handling ambiguity, the learning capability of the neural network, and the incentive design of the auction process; however, despite the practical benefits like increased demand forecasting accuracy and optimized resource allocation, the model faces limitations such as high dependence on the quality of historical sales data and the inherent computational complexity.

The critical need for accurate sales forecasting to enhance profitability in the retail sector provided the motivation for developing an improved solution, as existing deep learning (DL) models often exhibit limited real-world accuracy because they typically handle temporal or spatial patterns in isolation or omit critical contextual variables [10]. To overcome these limitations, the methodology proposed a novel hybrid CNN-LSTM model tailored for multivariate retail sales forecasting, wherein Convolutional Neural Networks (CNNs) efficiently capture short-term feature patterns and spatial dependencies, while Long Short-Term Memory (LSTM) networks model long-term temporal dependencies like seasonality and trends; crucially, the input was augmented with environmental factors (Effective Rain, Effective Temperature) and demographic factors (holidays, salary days, protests) using a publicly available real-world dataset. Evaluating the model against baseline approaches using real retail data demonstrated a significant improvement in predictive performance, yielding a high accuracy evidenced by a Mean Absolute Percentage Error (MAPE) of 4.16%, alongside a Mean Absolute Error (MAE) of 11,165 and a Root Mean Square Error (RMSE) of 26,200, confirming its superiority over standalone models. This high performance is attributed to the combined ability of the hybrid architecture to capture both short-term variations and long-term trends, especially since incorporating contextual variables significantly improved accuracy by allowing the model to anticipate demand shifts related to external factors like weather or events; nevertheless, a key limitation is that the dataset is restricted to a single store, thereby constraining the model's generalizability across diverse regions or retail chains, and internal analysis suggested that the model exhibited signs of over-fitting due to its success in memorizing dataset-specific sales spikes.

The recognized importance of accurate sales forecasting for optimizing critical business decisions, including inventory management, resource allocation, and financial planning, motivated the development of a high-performance deep learning solution designed specifically for low-dimensional tabular data [11]. The methodology proposed involved a computational framework based on streamlined transformers, drawing inspiration from the TabTransformer architecture but incorporating minor modifications, such as adding an attention layer after the normalization step for numerical features. The model processes categorical features through Column Embedding and Multi-head Attention layers, and numerical features through Normalization and subsequent Feed Forward layers that generate the Query, Key, and Value vectors for a Self-Attention layer; the resulting categorical and numerical attention outputs are then concatenated and fed into a Multi-layer Perceptron (MLP) for final prediction. When rigorously assessed against conventional machine learning models (k-NN, LR, RF, XGB) and other deep learning models (GRU, LSTM, TabTransformer baseline) over 15 iterations using the time-series data from the Kaggle "Walmart Recruiting-Store Sales Forecasting" competition, the transformer framework demonstrated superior stability and performance. The experimental results confirmed that the model achieved reduced prediction errors across metrics, including a mean Mean Absolute Error (MAE) of 3.12 and a mean Root Mean Square Error (RMSE) of 5.08, while yielding a consistently high R2 value of nearly 0.95. Although the enhanced accuracy and stability confirm the potential of this model for facilitating more informed decision-making in corporate operations, the study acknowledges key limitations: specifically, the evaluation was conducted on a single, specific medium-sized dataset, which restricts the applicability of the findings to similar topics and suggests the model may not fully capture the diversity and complexity of real-world problems.

The significant potential of crises, such as the COVID-19 pandemic, to profoundly disrupt supply chains, cause forecasting errors, and shift consumer demand necessitated the development of adaptable sales forecasting strategies that could account for these turbulent conditions, addressing the gap where most existing literature focuses only on normal operating environments [12]. To achieve this, the methodology employed several Machine Learning (ML) algorithms, including Gradient Boosting, CatBoost, and Multi-Layer Perceptron (MLP), to analyze consumer behavior and predict sales across six main product categories of a Turkish women's clothing retailer. This involved combining pre-crisis (2017-2020) and crisis (2020-2022) sales data, augmented with independent variables such as per capita and target audience population size, to better capture factors affecting sales. The results demonstrated that the performance of the ML algorithms varied significantly based on the product group: the Gradient Boosting algorithm achieved the highest value

of 94% in the top wear category, and along with CatBoost, excelled in product groups that experienced the most significant sales changes. Conversely, the MLP algorithm proved most effective in low-volume categories like shoes and accessories, achieving the lowest value of 0.20 in the shoe category. These findings confirm the efficacy of using ML models (specifically Gradient Boosting, CatBoost, and MLP) to adapt sales forecasts during crises, thereby enhancing retail resilience and optimizing strategic decisions like inventory management; however, a primary limitation is that the findings' generalizability is restricted as the study relied exclusively on data from a single women's clothing retailer in Turkiye and deliberately excluded online sales data from the analysis.

The necessity for accurate demand prediction in the highly volatile market of short-cycle agricultural products served as the impetus for developing an advanced forecasting solution, aiming to match supply with consumer needs and significantly reduce food waste [13]. To address the challenges of highly volatile and discontinuous data, the proposed methodology introduced a hierarchical prediction model combining Random Forest (RF) and Extreme Gradient Boosting (XGBoost). The first layer utilized RF to perform an initial prediction based on correlation features extracted via Grey Relation Analysis (GRA), which identified cooperative or competitive relationships among products, and subsequently extracted the residuals (errors) between predicted and actual values. In the second layer, XGBoost was employed, utilizing a new feature set derived from applying hierarchical clustering to classify the characteristics of the residuals, with the final prediction being the incorporation of the results from both layers. This approach demonstrated superior performance over standalone RF and XGBoost models when evaluated using three years of agricultural product sales data from a Chinese supermarket, achieving a Mean Absolute Percentage Error (MAPE) reduction of 10% and 12%, respectively, and a coefficient of determination () increase of 22% and 24%, respectively. While the results validate the model's capacity to enhance short-term forecasting precision and optimize the supply chain by combining the strengths of RF (robustness against nonlinear data) and XGBoost (good performance with structured data), a critical limitation is that the evaluation was performed on a single, specific medium-sized dataset which may restrict its applicability to the full diversity and complexity of real-world problems.

The imperative for accurate drug sales forecasting stems from its direct relationship to setting accurate target inventory, which helps businesses avoid shortages and inventory backlogs amidst the complicated and changeable market environment [14]. To enhance forecasting performance and overcome the inherent shortcomings of traditional Neural Networks (NNs) like long training times, slow convergence speed, and susceptibility to local minima, a novel Particle Swarm Optimization Neural Network (PSONN) model was proposed. This methodology involves using an improved Particle Swarm Optimization (PSO) algorithm to globally search and optimize the weights and thresholds of a Backpropagation Neural Network (BPNN), leveraging PSO's advantages in global optimization and fast convergence, and NN's ability to approximate arbitrary nonlinear functions. When tested using the 2019 sales data of a pharmaceutical company, the PSONN model demonstrated superior prediction effectiveness, achieving a high prediction accuracy of 96.14%, which is 12.78% higher than the traditional BP model. Furthermore, the PSONN model exhibited a lower error rate and shorter running time compared to AdaBoost-BP and AdaBoost-ELM models. While the results confirm that the model is practical and feasible for drug sales forecasting, simplifying the network structure, and improving convergence speed, a key limitation is the inherent randomness in PSONN and particle swarm improvement, which may lead to poor prediction stability. Additionally, the model comparison in the study did not include the testing of some other algorithms related to sales forecasting.

The need for accurate sales forecasting is crucial for the healthy development of enterprises in the increasingly competitive Cross-Border E-commerce (CBEC) market, where merchandise mix and prediction accuracy are essential prerequisites for market orientation and inventory control, especially for short life cycle products like fast fashion apparel, whose demand forecast is challenging [15]. To address this challenge and enhance prediction accuracy, a BP-LSTM-Verhulst nested neural network based on the AP algorithm was constructed. This method combines the advantages of three models: the BP neural network predicts sales by combining current independent variables; the LSTM neural network explores the influence of historical data on sales, leveraging the recurrent structure's memory function to consider the impact of earlier data on later predictions; and the Verhulst model predicts sales based on the growth trend of variables. The effectiveness of this nested neural network was demonstrated by an error comparison using a practical example, which proved that the combined model has higher accuracy in predicting the sales of goods compared to exponential regression and shallow neural networks [16][17]. While the deep learning prediction method is effective, especially by incorporating unstructured data like images (which are proven to be important marketing information online and significantly impact user purchasing behavior), and by demonstrating that combined models generally outperform single models, the study implicitly acknowledges limitations inherent in the BP neural network component, such as its lack of global search capability and susceptibility to falling into a local optimum

3. Method

In this study, we utilized the dataset provided in the Kaggle competition "Walmart Recruiting: Store Sales Forecasting". The dataset consists of three data frames that were merged to create a unified dataset, containing 14 variables divided into three categories [18]:

1) Continuous variables (11): Size, Temperature, Fuel_Price, CPI, Unemployment, Total_MarkDown, Max, Min, Mean, Median, and Std.
2) Categorical variables (3): Store, Dept, and Type.
3) Target variable: Weekly_Sales, used as the label for prediction tasks.

Given the time-series nature of the dataset, the data was split into three sets based on temporal ordering:

1) Training set: Comprising 331,742 samples recorded between 05/02/2010 and 30/03/2012 (113 weeks).
2) Validation set: Consisting of 44,192 samples from 06/04/2012 to 13/07/2012 (15 weeks).
3) Test set: Including 44,278 samples recorded between 20/07/2012 and 26/10/2012 (15 weeks).

The combined dataset spans 143 weeks (05/02/2010 to 26/10/2012) and comprises 420,212 samples, including both temporal and structured data. **Table 1** shows dataset information.

Table 1. Dataset Information

Data	Samples	Time Period	Weeks
Training	331,742	05/02/2010 - 30/03/2012	113
Validation	44,192	06/04/2012 - 13/07/2012	15
Test	44,278	20/07/2012 - 26/10/2012	15

This dataset provides a robust basis for evaluating the hybrid LSTM-MLP framework for sales forecasting, offering a mix of both temporal and numerical features crucial for model training and evaluation. This research employs a hybrid framework combining Long Short-Term Memory (LSTM) networks and Multi-Layer Perceptrons (MLPs) to leverage the strengths of both architectures for enhanced sales forecasting. Fig. 1 illustrates the proposed method of data flow and integration of LSTM and MLP components.

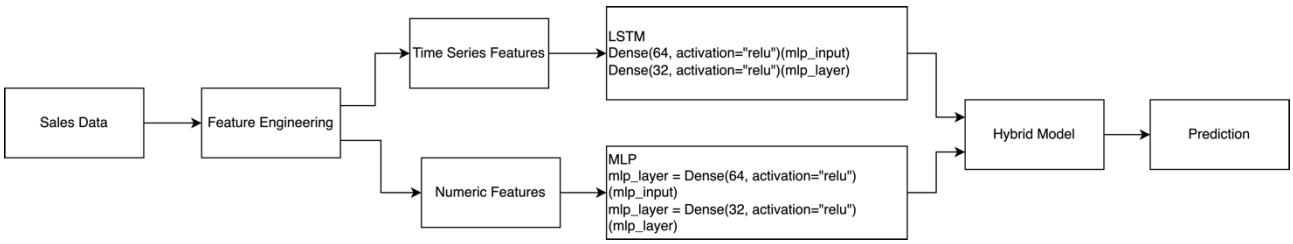


Fig. 1. Proposed Method: overview of the data flow and integration of LSTM and MLP components

According to **Fig. 1**, the methodology involves the following key steps:

1) Data Preprocessing
   - Handling Missing Values: Missing data points were either filled using appropriate imputation methods or dropped if they had negligible influence on the dataset.
   - Categorical Encoding: The categorical variables (e.g., Store, Dept, and Type) were converted into numerical representations using one-hot encoding.
   - Feature Scaling: Continuous variables were normalized to ensure compatibility with model training and to improve convergence.
   - Dataset Splitting: The data was split into training, validation, and test sets based on temporal ordering.
2) Feature Engineering
   - Time Series Features: Weekly sales patterns and other seasonal attributes were derived to capture temporal dependencies.
   - Statistical Features: Aggregated metrics such as rolling averages, trends, and deviations were computed to enhance the model's understanding of historical patterns.
3) Hybrid Model Architecture
   - LSTM:
     - Designed to capture temporal dependencies from time-series data.
     - Consists of two dense layers: one with 64 nodes and the second with 32 nodes, both utilizing the ReLU activation function.
   - MLP:
     - Processes non-sequential numerical features.
     - Features two dense layers with 64 and 32 nodes, employing ReLU activation
     - Integration and Prediction: Outputs from the LSTM and MLP components are concatenated, and the combined feature representation is passed through additional dense layers to produce the final prediction.

4) Evaluation Metrics

To comprehensively evaluate model performance, the following metrics were employed: Mean Absolute Error (MAE), Mean Squared Error (MSE) [19]-[21], Root Mean Squared Error (RMSE) [22], Mean Absolute Percentage Error (MAPE)[23], Coefficient of Determination (R^2) [25]-[31]. These steps collectively form the proposed methodology to improve the accuracy and efficiency of sales forecasting. The mathematical formula of these metrics expressed in Equation:

$$MAE = \frac{1}{N} \sum_{i=1}^{N} | yi - \gamma \hat{} i| \quad (1)$$

$$MSE = \frac{1}{N} \sum_{i=1}^{N} (yi - \gamma' i|)^2 \quad (2)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^{N} (yi - \gamma' i|)^2} \quad (3)$$

$$MAPE = \frac{1}{N} \sum_{i=1}^{N} | \frac{yi - \gamma' i}{yi} | \quad (4)$$

$$R^2 = \sum_{i=1}^{N} (\frac{yi - \gamma' i}{yi}) \quad (5)$$

3. Results and Discussion

3.1 Model Performance

The proposed hybrid LSTM-MLP model was evaluated using standard regression metrics, and the results were compared with state-of-the-art models. Our proposed model demonstrated superior performance across all metrics **Table 2** summarizes the results obtained from our model.

Table 2. Performance Metrics and Result of Proposed Method

Metric	Value
MAE	5.4819e-05
MSE	6.6741e-08
RMSE	0.000258
MAPE	16.444e-06
R^2	0.999937

3.2 Comparative Analysis

To provide a meaningful comparison, **Table 3** outlines the performance of our model against other baseline methods from the main reference. The proposed model consistently outperformed all other approaches.

Table 3. Performance Comparison with Baseline Models

Model	MSE	RMSE	MAE	MAPE	R^2
k-Nearest Neighbor	213.71	14.62	8.68	72.53	0.55
Linear Regression	434.88	20.85	14.64	166.71	0.09
Random Forest	64.56	8.03	5.27	60.99	0.87
eXtreme Gradient Boosting	239.44	15.47	9.61	102.68	0.5
GRU-based Model	37.84	6.15	3.76	16.08	0.92
LSTM-based Model	34.64	5.89	3.56	30.26	0.93
TabTransformer	85.72	9.26	6.21	79.89	0.82
Transformer Based MLP [1]	25.76	5.08	3.12	9.49	0.95
Proposed scheme	**6.67**	**0.0002**	**5.48**	**16.44**	**0.99**

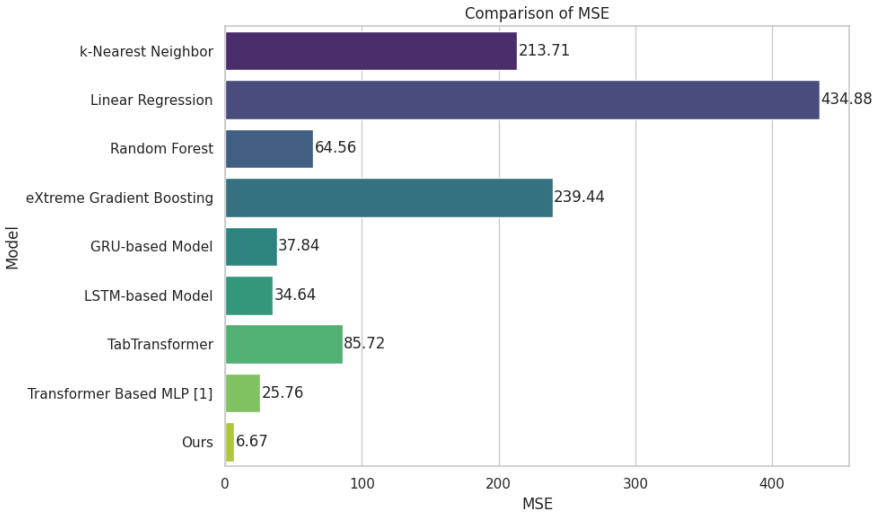


Fig. 2. Visualizes the performance of our model against other baseline methods from the main reference.

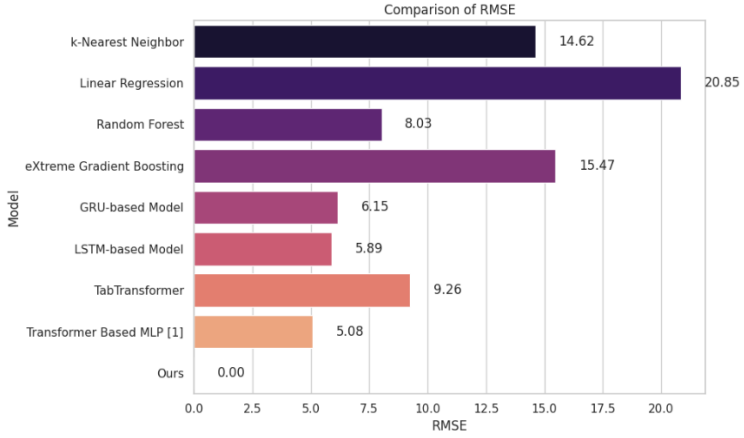


Fig. 3. RMSE Comparison

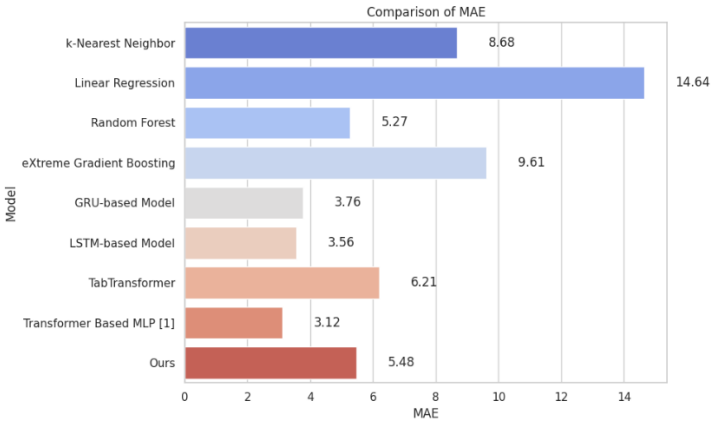


Fig. 4. MAE Comparison

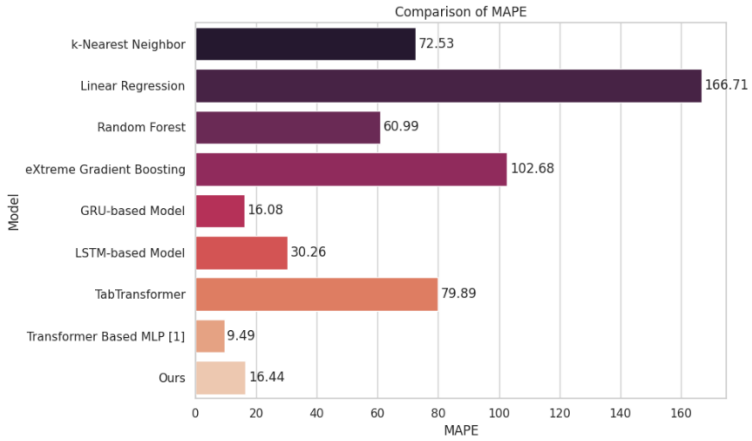


Fig. 5. MAPE Comparison

The results visualizations is shown in Figs. 2, 3, 4, and 5. The proposed Hybrid LSTM-MLP framework has showcased remarkable performance compared to traditional models and recent state-of-the-art methods. The inclusion of an LSTM component efficiently captured temporal patterns in the data, while the MLP handled numerical features, enhancing overall performance. The hybrid model offers flexibility and scalability for other time-series datasets with mixed feature types, extending its usability to domains beyond sales forecasting. The model demonstrates computational efficiency, making it suitable for real-world applications, especially those requiring accurate predictions under resource constraints.

4. Conclusion

In this study, we proposed a hybrid framework combining Long Short-Term Memory (LSTM) networks and Multi-Layer Perceptrons (MLP) to enhance sales forecasting through effective feature processing. The LSTM component efficiently captured temporal dependencies in time-series data, while the MLP handled numerical features, enabling a comprehensive integration of various data types. The proposed model was rigorously evaluated using the Walmart Store Sales Forecasting dataset, demonstrating superior performance in terms of MSE, RMSE, MAE, and R^2 compared to traditional machine learning models and state-of-the-art methods. Specifically, the model achieved remarkable results with an RMSE of 0.000258, MAE of 5.48, and R^2 of 0.9999, highlighting its predictive accuracy and stability. These outcomes surpass conventional approaches, showing the advantage of leveraging hybrid neural architecture for sales prediction tasks. The results underline the efficacy of combining time-series and numerical feature learning within a single framework, offering significant contributions to forecasting tasks involving tabular data. Furthermore, lightweight architecture and efficient computation make it suitable for resource-constrained environments, such as embedded systems or IoT devices. This work lays a foundation for future research in hybrid neural network frameworks and feature engineering. Future studies could explore broader datasets, incorporate advanced neural components like transformers, or integrate domain-specific knowledge to further enhance performance and applicability in real-world scenarios.

Conflicts of Interest

The author declares no conflict of interest.

References

[1] Y. Sun and T. Li, "A transformer-based framework for enterprise sales forecasting," *PeerJ Computer Science*, vol. 10, pp. 1–14, 2024, doi: 10.7717/peerj-cs.2503.

[2] R. Murugesan, E. Mishra, and A. H. Krishnan, "Forecasting agricultural commodities prices using deep learning-based models: basic LSTM, bi-LSTM, stacked LSTM, CNN LSTM, and convolutional LSTM," *Int. J. Sustain. Agric. Manag. Inform.*, vol. 8, no. 3, p. 242, 2022, doi: 10.1504/IJSAMI.2022.125757.

[3] H. Li and D. Wei, "Prediction based on traditional network prediction model and LSTM deep neural network model," in *Proc. 4th Int. Conf. Comput., Netw. Internet Things*, New York, NY, USA: ACM, May 2023, pp. 726–730, doi: 10.1145/3603781.3603910.

[4] K. Wang, H. Du, R. Jia, and H. Jia, "Performance comparison of Bayesian deep learning model and traditional Bayesian neural network in short-term PV interval prediction," *Sustainability (Switzerland)*, vol. 14, no. 19, Oct. 2022, doi: 10.3390/su141912683.

[5] S. Gyamerah and D. R. Korda, "Prediction of stock market returns using LSTM model and traditional statistical model," 2021.

[6] D. Brykin, "Sales forecasting models: comparison between ARIMA, LSTM and Prophet," *J. Comput. Sci.*, vol. 20, no. 10, pp. 1222–1230, 2024, doi: 10.3844/jcssp.2024.1222.1230.
[7] N. Qassrawi, M. Azzeh, and M. Hijjawi, "Drug sales forecasting in the pharmaceutical market using deep neural network algorithms," *Int. J. Syst. Innov.*, vol. 8, no. 3, pp. 63–83, 2024, doi: 10.6977/IJoSI.202409_8(3).0006.
[8] I. Sonata and Y. Heryadi, "Comparison of LSTM and transformer for time series data forecasting," in *Proc. 7th Int. Conf. Informat. Comput. Sci. (ICICoS)*, IEEE, Jul. 2024, pp. 491–495, doi: 10.1109/ICICoS62600.2024.10636892.
[9] Z. Gong, "Optimization of cross-border E-commerce (CBEC) supply chain management based on fuzzy logic and auction theory," *Sci. Rep.*, vol. 14, p. 14088, 2024.
[10] S. Mansur, K. Sattar, S. E. Hosseini, S. Pervez, I. Ahmad, K. Saleem, and A. Z. Elhendi, "Sales forecasting for retail stores using hybrid neural networks and sales-affecting variables," *PeerJ Comput. Sci.*, vol. 11, e3058, 2025, doi: 10.7717/peerj-cs.3058.
[11] Y. Sun and T. Li, "A transformer-based framework for enterprise sales forecasting," *PeerJ Comput. Sci.*, vol. 10, e2503, 2024, doi: 10.7717/peerj-cs.2503.
[12] K. T. Kizgin, S. Alp, N. Aydin, and H. Yu, "Machine learning-based sales forecasting during crises: Evidence from a Turkish women's clothing retailer," *Sci. Prog.*, vol. 108, no. 1, pp. 1–18, 2025.
[13] J. Li, B. Lin, P. Wang, Y. Chen, X. Zeng, X. Liu, and R. A. Chen, "A hierarchical RF-XGBoost model for short-cycle agricultural product sales forecasting," *Foods*, vol. 13, p. 2936, 2024, doi: 10.3390/foods13293936.
[14] C. Yu, "Design of drug sales forecasting model using particle swarm optimization neural networks model," *Comput. Intell. Neurosci.*, vol. 2022, Art. ID 6836524, 9 pp., doi: 10.1155/2022/6836524.
[15] L. Tian and X. Wang, "A dynamic prediction neural network model of cross-border e-commerce sales for virtual community knowledge sharing," *Comput. Intell. Neurosci.*, vol. 2022, Art. ID 2529372, 11 pp., doi: 10.1155/2022/2529372.
[16] Q. Li and M. Yu, "Achieving sales forecasting with higher accuracy and efficiency: A new model based on modified transformer," *J. Theor. Appl. Electron. Commer. Res.*, vol. 18, no. 4, pp. 1990–2006, Dec. 2023, doi: 10.3390/jtaer18040100.
[17] G. Woo, C. Liu, D. Sahoo, A. Kumar, and S. Hoi, "ETSformer: Exponential smoothing transformers for time-series forecasting," *arXiv preprint* arXiv:2202.01381, Feb. 2022.
[18] H. M. Merdas and A. H. Mousa, "Food sales prediction using MLP, RANSAC, and bagging," *J. Tech.*, vol. 5, no. 4, pp. 202–208, Dec. 2023, doi: 10.51173/jt.v5i4.1458.
[19] I. Vallés-Pérez, E. Soria-Olivas, M. Martínez-Sober, A. J. Serrano-López, J. Gómez-Sanchís, and F. Mateo, "Approaching sales forecasting using recurrent neural networks and transformers," *Expert Syst. Appl.*, vol. 201, Sep. 2022, doi: 10.1016/j.eswa.2022.116993.
[20] Y. Li, Y. Yang, K. Zhu, and J. Zhang, "Clothing sale forecasting by a composite GRU–Prophet model with an attention mechanism," *IEEE Trans. Ind. Inform.*, vol. 17, no. 12, pp. 8335–8344, 2021.
[21] S. Beheshti-Kashi, H. R. Karimi, K.-D. Thoben, M. Lütjen, and M. Teucke, "A survey on retail sales forecasting and prediction in fashion markets," *Syst. Sci. Control Eng.*, vol. 3, no. 1, pp. 154–161, 2014, doi: 10.1080/21642583.2014.999389.
[22] V. Sohrabpour, P. Oghazi, R. Toorajipour, and A. Nazarpour, "Export sales forecasting using artificial intelligence," *Technol. Forecast. Soc. Change*, vol. 163, p. 120480, 2021, doi: 10.1016/j.techfore.2020.120480.
[23] R. K. Paul, M. Yeasin, P. Kumar, P. Kumar, M. Balasubramanian, H. S. Roy, A. K. Paul, and A. K. Gupta, "Machine learning techniques for forecasting agricultural prices: A case of brinjal in Odisha, India," *PLoS ONE*, vol. 17, e0270553, 2022.
[24] S. A. Haider, S. R. Naqvi, T. Akram, G. A. Umar, A. Shahzad, M. R. Sial, S. Khaliq, and M. Kamran, "LSTM neural network based forecasting model for wheat production in Pakistan," *Agronomy*, vol. 9, p. 72, 2019.
[25] Z. Liu, "Forecasting stock prices based on multivariable fuzzy time series," *AIMS Math.*, vol. 8, pp. 12778–12792, 2023.
[26] S. M. Gonzales, H. Iftikhar, and J. L. López-Gonzales, "Analysis and forecasting of electricity prices using an improved time series ensemble approach: An application to the Peruvian electricity market," *AIMS Math.*, vol. 9, pp. 21952–21971, 2024.
[27] D. Wu, B. Lu, and Z. Xu, "Price forecasting of marine fish based on weight allocation intelligent combinatorial modelling," *Foods*, vol. 13, p. 1202, 2024.
[28] M. Yucesan, M. Gül, and E. CelikGul, "Performance comparison between ARIMAX, ANN and ARIMAX-ANN hybridization in sales forecasting for furniture industry," *Drvna Ind.*, vol. 69, no. 4, pp. 357–370, 2018.
[29] H. Lian, B. Liu, and P. Li, "A fuel sales forecast method based on variational Bayesian structural time series," *J. High Speed Netw.*, vol. 27, no. 1, pp. 1–22, 2021.
[30] Z. Duan, B. Gutierrez, and L. Wang, "Forecasting plug-in electric vehicle sales and the diurnal recharging load curve," *IEEE Trans. Smart Grid*, vol. 5, no. 1, pp. 527–535, 2014.
[31] A. Russo, P. G. Lind, and F. Raische, "Neural network forecast of daily pollution concentration using optimal meteorological data at synoptic and local scales," *Atmos. Pollut. Res.*, vol. 6, no. 3, pp. 540–549, 2015.