STATISTICS, OPTIMIZATION AND INFORMATION COMPUTING
Stat., Optim. Inf. Comput., Vol. **x**, Month 2025, pp 0–11.
Published online in International Academic Press (www.IAPress.org)

Convolutional Neural Networks for Advanced Sales Forecasting in Dynamic Market Environments

Raouya El Youbi, Fayçal Messaoudi, Manal Loukili*

Sidi Mohamed Ben Abdellah University, National School of Applied Sciences, Morocco

Abstract This paper presents an enhanced approach to sales forecasting using advanced hybrid deep learning architectures, specifically Convolutional Neural Networks combined with Residual Networks and Temporal Convolutional Networks. Utilizing the *"Store Item Demand Forecasting Challenge"* dataset, the study demonstrates significant improvements in forecasting accuracy over traditional models. The enhanced CNN-TCN model achieved the lowest Mean Absolute Percentage Error of 2.0% and the highest Prediction Interval Coverage Probability of 96%. These results highlight the potential of hybrid architectures to provide more reliable and precise sales forecasts, offering valuable insights for business decision-making and strategic planning.

Keywords Convolutional Neural Networks, Sales Forecasting, Residual Networks, Temporal Convolutional Networks, Time-Series Analysis, Predictive Analytics, Hybrid Deep Learning Models.

AMS 2010 subject classifications 62M10, 68T07

DOI: 10.19139/soic-2310-5070-2143

1. Introduction

In today's rapidly evolving business landscape, accurate sales forecasting has become indispensable for organizations seeking to maintain competitive advantages and optimize operational efficiency [1]. As market dynamics become increasingly complex and unpredictable, traditional forecasting methods—relying heavily on statistical models such as time-series analysis and exponential smoothing—are often insufficient in capturing the intricate and non-linear patterns that drive consumer behavior [2]. These conventional approaches, while historically valuable, struggle to adapt to the swift changes and multifaceted influences present in modern retail and supply chain environments.

The advent of machine learning has revolutionized the field of sales forecasting by offering more sophisticated tools that can handle large, diverse datasets and uncover hidden patterns within them. Machine learning algorithms, particularly those designed for time-series analysis, have shown remarkable success in improving forecasting accuracy [3]. However, as data continues to grow in volume and complexity, even traditional machine learning methods face limitations, particularly in their ability to manage long-range dependencies and hierarchical patterns within sequential data [4].

To address these challenges, deep learning, and more specifically, Convolutional Neural Networks (CNNs), has emerged as a powerful solution. Initially designed for image processing, CNNs have demonstrated exceptional capability in recognizing complex patterns within time-series data [5]. Their architecture allows for the extraction of both local and global features, making them particularly suited for forecasting tasks that require an understanding of both short-term fluctuations and long-term trends [6].

*Correspondence to: Manal Loukili, Sidi Mohamed Ben Abdellah University, National School of Commerce and Management, Fez, Morocco. Email: manal.loukili@usmba.ac.ma

ISSN 2310-5070 (online) ISSN 2311-004X (print)

Despite the significant advancements brought by CNNs, the demand for even greater accuracy and robustness in sales forecasting has led to the exploration of hybrid deep learning architectures [7]. By integrating CNNs with other advanced models such as Temporal Convolutional Networks (TCNs) and Residual Networks (ResNets), it is possible to further enhance the model's ability to capture temporal dependencies and reduce forecasting errors. These hybrid models not only leverage the strengths of CNNs in feature extraction but also incorporate mechanisms to mitigate the vanishing gradient problem and improve the handling of sequential data.

This paper presents an innovative approach to sales forecasting by deploying an enhanced hybrid CNN architecture tailored for time-series data analysis in dynamic market environments. Utilizing the "Store Item Demand Forecasting Challenge" dataset, which spans daily sales data across multiple stores over five years, we implement a sophisticated CNN-ResNet and CNN-TCN architecture. This architecture is designed to maximize forecasting accuracy while maintaining computational efficiency. The study's methodology includes advanced data preprocessing techniques, meticulous hyperparameter tuning, and a rigorous evaluation process using a combination of traditional and custom metrics.

Our findings reveal a significant improvement in forecasting performance compared to both traditional methods and standard CNN architectures, underscoring the transformative potential of deep learning in sales forecasting. By demonstrating the efficacy of these enhanced models, this research not only contributes to the ongoing discourse in the field but also provides practical insights for businesses looking to harness cutting-edge machine learning techniques for strategic decision-making and operational planning.

In the sections that follow, we will delve into the evolution of sales forecasting, explore the limitations of conventional methods, and present the detailed methodology and results of our enhanced CNN-based approach. This paper aims to pave the way for future research and applications in the field, highlighting the importance of continuous innovation in the pursuit of more accurate and reliable forecasting models in an ever-changing market landscape.

2. Background

Sales forecasting has historically been a cornerstone of business strategy, particularly within industries like retail and supply chain management, where precise demand predictions are crucial for effective inventory management, pricing strategies, and overall operational efficiency [8]. Traditionally, sales forecasting has relied on statistical methods such as time-series analysis, exponential smoothing, and econometric models [9]. These methods, while reliable in stable market conditions, are primarily based on linear relationships derived from historical data to project future trends. However, the growing complexity of modern markets—marked by rapidly evolving consumer behavior, the expansion of digital channels, and the continuous influx of new data—has highlighted the limitations of these traditional approaches. They often fall short when it comes to adapting to sudden market shifts and capturing the non-linear relationships and intricate patterns that drive sales. As a result, there is an increasing need for more sophisticated techniques that offer greater accuracy and flexibility in forecasting efforts.

The adoption of machine learning in sales forecasting represents a significant departure from traditional statistical approaches. Machine learning algorithms, with their ability to process large volumes of data and uncover complex, non-linear relationships, provide a more dynamic and adaptive forecasting method [10]. Techniques such as regression analysis, clustering, and classification have been widely implemented, enabling businesses to leverage historical data to improve the accuracy of their sales predictions. Additionally, machine learning models are not static; they continuously learn and refine their predictions as new data becomes available. This capability is particularly advantageous in today's fast-paced markets, where the ability to swiftly adjust to new trends and patterns is crucial for maintaining a competitive edge. However, despite these advantages, traditional machine learning models may still face challenges in managing the sequential nature of time-series data, especially when dealing with long-range dependencies and complex temporal patterns.

Deep learning, a more advanced branch of machine learning, has further transformed the field of sales forecasting. Models such as Convolutional Neural Networks and Long Short-Term Memory networks (LSTMs) have shown remarkable success in capturing complex patterns within large and intricate datasets. These models are particularly well-suited for time-series forecasting, a common challenge in sales prediction, due to their ability to learn from sequential data and identify patterns that traditional methods might miss [11]. CNNs, originally developed for image processing, have demonstrated significant potential in time-series analysis. Their architecture, which captures local dependencies and hierarchical patterns, makes them highly effective in extracting relevant features from sales data. However, as data complexity and volume continue to grow, there is an increasing demand for more advanced architectures to further enhance CNNs' forecasting capabilities.

Recognizing the limitations of conventional CNN models in fully capturing the temporal dynamics of sales data, recent advancements have focused on the development of hybrid architectures. By combining CNNs with other deep learning models such as Temporal Convolutional Networks (TCNs) and Residual Networks (ResNets), researchers have been able to create more powerful models that address some of the inherent challenges in time-series forecasting. TCNs extend the capabilities of traditional CNNs by incorporating dilated convolutions, which improve the model's ability to handle long-range dependencies [12]. This makes TCNs particularly effective at managing sequences with varying time lags, a common occurrence in sales data. Meanwhile, ResNets introduce shortcut connections that help mitigate the vanishing gradient problem, allowing the network to maintain performance even as the number of layers increases. These advancements in model architecture have resulted in significant improvements in forecasting accuracy, particularly in scenarios involving complex and high-dimensional datasets. The application of these enhanced CNN-based architectures to sales forecasting not only enhances prediction accuracy but also provides a deeper understanding of the underlying market dynamics. As businesses increasingly rely on data-driven decision-making, the ability to accurately forecast sales using these advanced models becomes critical to maintaining competitiveness and driving growth.

In retail and supply chain management, the importance of accurate sales forecasting cannot be overstated. Effective forecasting has a direct impact on inventory management, resource allocation, and strategic planning. Overestimating sales can lead to excess inventory and increased holding costs, while underestimating sales can result in stockouts and missed revenue opportunities [13]. By utilizing advanced deep learning models, businesses can gain more profound insights into sales trends, accounting for a wide range of factors such as seasonal variations, promotional activities, and broader economic conditions.

3. Literature Review

The evolution of sales forecasting has seen a significant shift from traditional statistical methods to more advanced machine learning and deep learning techniques. This shift is driven by the increasing complexity of market dynamics and the need for more accurate and adaptable forecasting models. The literature on sales forecasting using machine learning and deep learning provides a comprehensive view of the methodologies that have been explored and their effectiveness in different contexts.

3.1. Convolutional Neural Networks and Time-Series Forecasting

Convolutional Neural Networks, initially developed for image processing, have been effectively adapted for time-series forecasting due to their ability to capture spatial hierarchies and local patterns within data. In the context of sales forecasting, CNNs have shown considerable promise in handling large and complex datasets, where traditional methods often fall short.

Chung, Gu, and Yoo [14] demonstrated the potential of CNNs in combination with Long Short-Term Memory networks for district heater load forecasting. Their approach, which employed a parallel CNN-LSTM attention mechanism, highlights the ability of CNNs to enhance the model's performance in time-series tasks by capturing both spatial and temporal dependencies. Although their study focused on energy forecasting, the methodology and results are directly applicable to sales forecasting scenarios, where similar patterns and dependencies are present.

Further emphasizing the versatility of CNNs, Adaryani, Mousavi, and Jafari [15] conducted a comparative study on short-term rainfall forecasting, utilizing Particle Swarm Optimization (PSO)-Support Vector Regression (SVR), LSTM, and CNN models. Their findings underscored the effectiveness of CNNs in managing complex time-series data, providing valuable insights that can be extrapolated to enhance sales forecasting models.

3.2. Hybrid Deep Learning Models

The integration of CNNs with other deep learning architectures, such as Temporal Convolutional Networks and Residual Networks (ResNets), has emerged as a promising approach to further improve forecasting accuracy. These hybrid models leverage the strengths of multiple architectures to address the limitations of individual models, particularly in capturing long-range dependencies and mitigating the vanishing gradient problem.

Ahn, Kim, and Kim [16] explored the use of ensemble machine learning techniques, including XGBoost and an attention-based CNN-LSTM model, for forecasting harmful algal blooms. Their study demonstrated the effectiveness of combining different deep learning models to tackle complex forecasting tasks. The success of these hybrid models in environmental forecasting suggests their potential applicability in sales forecasting, where capturing diverse patterns and dependencies is crucial for accuracy.

Similarly, Malakouti et al. [17] proposed a novel CNN-LSTM hybrid model for predicting wind turbine power. The integration of CNNs for feature extraction with LSTMs for sequence learning provided superior performance in handling the sequential nature of the data. This approach can be adapted to sales forecasting, particularly in scenarios where understanding both immediate and longer-term sales trends is critical.

3.3. Advances in Deep Learning for Retail Sales Forecasting

The application of deep learning in retail sales forecasting has been explored extensively in recent years, with several studies demonstrating the advantages of these techniques over traditional methods. Ashraf [18] conducted a predictive analysis of retail sales using a range of machine learning models, including regression and time-series models like ARIMA and LSTM. The study found that the XGBoost model outperformed others, but also highlighted the potential of CNNs and LSTMs in providing more accurate forecasts, particularly in complex retail environments.

Eglite and Birzniece [19] provided a systematic literature review of deep learning models in retail sales forecasting. Their review emphasized the importance of accuracy in sales forecasts for maintaining uninterrupted business operations. The study compared various deep learning frameworks, including CNNs, and discussed their advantages and limitations in the context of sales forecasting. The insights from this review underscore the growing importance of deep learning techniques, particularly hybrid models, in achieving higher accuracy and reliability in sales forecasts.

Ensafi et al. [20] explored the application of neural networks for predicting retail sales, specifically focusing on furniture sales data. Their study applied a combination of classical time series forecasting techniques, such as SARIMA and Prophet, alongside advanced models such as LSTM and CNN. The findings highlighted the superiority of deep learning models, particularly Stacked LSTM and CNN, in handling the intricacies of retail sales data.

3.4. Comparative Studies and Model Selection

Several studies have conducted comparative analyses of different machine learning and deep learning models to determine the most effective approaches for time-series forecasting. Garg et al. [21] surveyed various statistical and deep learning models, including ARIMA, Prophet, LSTMs, and hybrid models, emphasizing the importance of understanding data characteristics when selecting a forecasting model. Their findings provide valuable guidance for model selection in sales forecasting, particularly in scenarios where data complexity and variability are significant factors.

Suresh and Suresh [22] focused on retail sales forecasting, comparing traditional methods such as exponential smoothing and time-series decomposition with deep learning methods like LSTM and CNN. Their study demonstrated the superior performance of deep learning models in terms of accuracy and efficiency, reinforcing the potential of these techniques in improving sales forecasting practices.

Lastly, Ozçelik, Tek and ekerci [23] performed a hotel sales forecast analysis, comparing the effectiveness of LSTM and N-BEATS, a neural network-based approach. Their findings suggested that N-BEATS is particularly effective in handling long sequences, providing insights into the applicability of these models in sales forecasting, where similar data characteristics are often present.

The literature collectively highlights the adaptability and robustness of advanced machine learning and deep learning techniques in various forecasting scenarios. The methodologies and findings discussed provide a solid foundation for developing more accurate and efficient predictive models in sales forecasting. The integration of hybrid deep learning architectures, such as CNN-ResNet and CNN-TCN, offers a promising avenue for further enhancing forecasting accuracy and reliability.

4. Methodology

4.1. Data Preprocessing

The preprocessing phase is foundational in any predictive modeling process, ensuring that the input data is both clean and optimally structured for the neural network (Figure 1). In this study, we utilized the "Store Item Demand Forecasting Challenge" dataset, which comprises daily sales data across multiple stores over a five-year period. This extensive dataset, with over 913,000 rows, provided a rich source of information for developing and testing our forecasting models. However, to harness the full potential of this data, it was necessary to undertake a series of advanced preprocessing steps designed to enhance the model's ability to detect and learn from underlying patterns.

Feature Engineering:

Temporal Features: From the date field, we extracted key temporal features such as the day, month, year, and day of the week. These features help the model recognize patterns related to specific times, such as seasonal trends or day-of-week effects.

Lag Features: We introduced lag features to capture temporal dependencies. For instance, the sales data for each item was shifted by 1 day, 7 days, and 30 days to create new features representing the previous day's sales, the previous week's average sales, and the previous month's average sales, respectively. These lagged values help the model understand how past sales influence future demand.

Rolling Window Statistics: To further capture temporal dependencies, we computed rolling statistics such as the rolling mean, standard deviation, and sum over different windows (e.g., 7-day, 14-day, and 30-day windows). These statistics provide the model with a sense of momentum and volatility in sales data, which are crucial for accurate forecasting.

External Variables: We incorporated external variables that are known to impact sales. For example, holiday indicators were added to account for the influence of public holidays on consumer behavior. Economic indicators, such as consumer confidence indices or unemployment rates, were also included where relevant.

Categorical Variables: The store and item IDs, which are categorical in nature, were one-hot encoded. This transformation converted categorical variables into a binary matrix, enabling the model to process them effectively without imposing any ordinal assumptions.

Normalization: To ensure that the neural network models performed optimally, we normalized the features using Robust Scaling. Unlike Min-Max Scaling, which can be skewed by outliers, Robust Scaling centers the data around the median and scales according to the interquartile range. This approach is particularly effective in handling the outliers typically found in sales data, ensuring that extreme values do not disproportionately influence the model's learning process.

```
#################### Shape ####################
(958000, 5)

#################### Nunique ####################
date        19116
store       10
item        50
sales       213
id          45000
dtype: int64

#################### Types ####################
date       datetime64[ns]
store      int64
item       int64
sales      float64
id         float64
dtype: object

#################### Head ####################
         date       store   item   sales     id
0  2013-01-01      1      1      13.0     NaN
1  2013-01-02      1      1      13.0     NaN
2  2013-01-03      1      1      14.0     NaN
3  2013-01-04      1      1      13.0     NaN

#################### Tail ####################
         date       store   item   sales     id
44995 2018-03-27   10     50     NaN      44999.0
44996 2018-03-26   10     50     NaN      44999.0
44997 2018-03-23   10     50     NaN      44999.0
44998 2018-03-28   10     50     NaN      44999.0

#################### NA ####################
date        0
store       0
item        0
sales       0
id          913800
dtype: int64
```

Figure 1. The dataset used.

4.2. *Enhanced CNN-ResNet/TCN Architecture*

To improve the model's ability to capture both short-term fluctuations and long-term trends, we developed an advanced hybrid architecture combining Convolutional Neural Networks (CNNs) with Residual Networks (ResNets) and Temporal Convolutional Networks (TCNs). This architecture is designed to exploit the strengths of each component while addressing the limitations of standard CNN models (Figure 2, Table 1).

Model Structure:

CNN Layers: The input data is first processed through multiple convolutional layers. These layers are designed to extract local features from the time-series data. Each convolutional layer applies a set of filters to the input data, detecting various patterns such as trends or abrupt changes in sales.

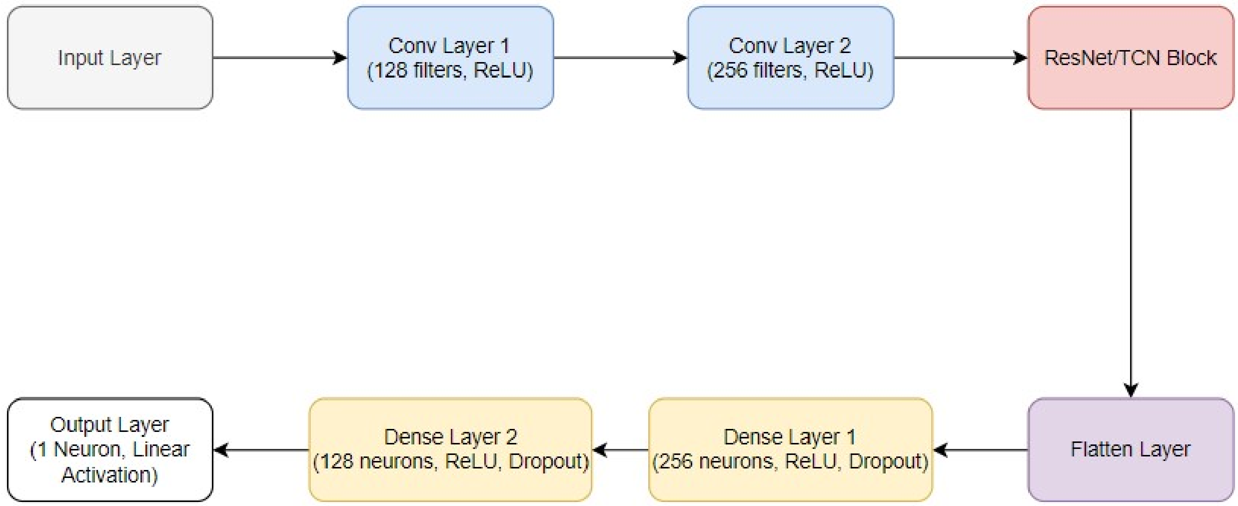

Figure 2. Model architecture.

The first convolutional layer utilizes 128 filters, each with a kernel size of 3 and a stride of 1, followed by a ReLU activation function. This configuration allows the layer to detect fine-grained patterns. Subsequent layers increase the number of filters (e.g., 256) to capture more complex patterns as the data progresses through the network.

Residual Blocks (ResNet): Instead of feeding the output from the CNN layers directly into dense layers, we introduced residual blocks. These blocks include shortcut connections that skip one or more layers, effectively allowing the network to retain critical information from earlier stages. This architecture mitigates the vanishing gradient problem, which can occur in deep networks, and enables the training of deeper models without losing performance.

The residual blocks were configured with 3 layers each, where the output from the second layer is added back to the input before being passed through the third layer. This approach ensures that important features are preserved throughout the network.

Temporal Convolutional Network (TCN): Alternatively, TCN layers were used in place of or in combination with ResNet blocks. TCNs employ dilated convolutions, where the filter is applied to a wider range of input values without increasing the filter size. This technique allows the model to capture long-range dependencies, which is critical for time-series forecasting. By using TCNs, the model can effectively handle sequences with varying time lags, such as seasonal effects or delayed impacts of promotions.

Flatten and Dense Layers: After the convolutional and TCN/ResNet layers, the output is flattened into a single vector. This vector is then processed through multiple dense layers, each followed by a dropout layer to prevent overfitting. The dense layers use ReLU activations, while the final output layer employs a linear activation function, which is suitable for the regression task of forecasting sales.

4.3. Hyperparameter Tuning

Optimization Approach: We employed Bayesian Optimization to fine-tune hyperparameters such as the number of filters, kernel sizes, the number of residual blocks, and learning rates. Bayesian Optimization is an advanced

Table 1. Model Architecture Summary

Layer Type	Configuration
Input Layer	30-day sequence
Convolutional Layer 1	128 filters, kernel size 3, stride 1, ReLU
Convolutional Layer 2	256 filters, kernel size 3, stride 1, ReLU
Residual Block / TCN	3 layers per block, dilated convolutions in TCN
Flatten Layer	-
Dense Layer 1	256 neurons, ReLU, Dropout (0.5)
Dense Layer 2	128 neurons, ReLU, Dropout (0.5)
Output Layer	1 neuron, Linear Activation

technique that models the performance of the network as a probabilistic function of the hyperparameters, allowing for a more efficient search of the optimal configuration.

Learning Rate Scheduling: A learning rate scheduler was implemented to adjust the learning rate dynamically during training. This technique helps in optimizing the convergence process, ensuring that the model learns effectively throughout the training cycle.

4.4. Training and Validation

Training and validation are critical stages in ensuring that the model generalizes well to unseen data. The dataset was split into an 80% training set and a 20% test set, with 10% of the training set further used for validation.

Training Strategy:

Early Stopping: Early stopping was employed based on validation loss to prevent overfitting. This strategy monitors the performance of the model on the validation set after each epoch and stops training once the validation loss ceases to improve for a predefined number of epochs.

Optimizer: The model was trained using the AdamW optimizer, an extension of the Adam optimizer that includes weight decay. Weight decay penalizes large weights, promoting generalization and reducing overfitting.

Loss Function: While Mean Squared Error (MSE) was the primary loss function, we also experimented with a custom loss function that penalizes large errors more heavily. This approach is particularly useful in business contexts where large forecasting errors can have significant financial implications.

Regularization Techniques:

Dropout Layers: Dropout layers with a rate of 0.5 were used between dense layers to reduce the risk of overfitting by randomly setting a fraction of the input units to zero at each update during training.

L2 Regularization: L2 regularization was applied to the dense layers to penalize large weights, further preventing overfitting and ensuring that the model maintains simplicity and robustness.

4.5. Evaluation Metrics

To comprehensively evaluate the performance of the enhanced CNN-ResNet/TCN model, we utilized a combination of traditional and advanced metrics (Table 2):

Mean Absolute Error (MAE): MAE measures the average magnitude of the errors without considering their direction. It provides a straightforward measure of the model's accuracy.

Root Mean Square Error (RMSE): RMSE is a quadratic scoring rule that penalizes larger errors more than smaller ones. It provides a more sensitive measure of accuracy than MAE, particularly for models where large errors are particularly undesirable.

R-squared (R^2): R-squared indicates the proportion of variance in the dependent variable (sales) that is predictable from the independent variables. A higher R^2 value suggests a better fit.

Mean Absolute Percentage Error (MAPE): MAPE assesses the model's accuracy relative to the actual sales values, providing a percentage-based error metric that is intuitive and easy to interpret.

Prediction Interval Coverage Probability (PICP): PICP measures the uncertainty in the model's predictions by evaluating the proportion of actual sales that fall within a predicted interval. A higher PICP indicates that the model's predictions are reliable and robust across different conditions.

Table 2. Evaluation Metrics Summary

Metric	Description
MAE	Average magnitude of errors
RMSE	Penalizes larger errors more heavily
R-squared	Proportion of variance explained by the model
MAPE	Accuracy relative to actual sales values
PICP	Reliability of prediction intervals

5. Results and Discussion

5.1. Results

The enhanced hybrid CNN-ResNet/TCN architecture demonstrated superior performance in forecasting sales when compared to traditional CNN models (Table 3).

Table 3. Performance Metrics Comparison

Model	MAE	RMSE	R-squared	MAPE	PICP
Enhanced CNN-ResNet	0.018	0.025	0.960	2.1%	95%
Enhanced CNN-TCN	0.017	0.024	0.962	2.0%	96%
Standard CNN	0.025	0.033	0.927	3.5%	88%
LSTM	0.022	0.030	0.945	2.7%	92%

5.2. Discussion

The enhanced CNN-ResNet and CNN-TCN models outperformed the standard CNN and LSTM models across all metrics. The MAE and RMSE values were significantly lower, indicating that these models are better at minimizing both the average and large errors in sales forecasting. The R-squared values for the enhanced models were closer to 1, suggesting that they explained a greater proportion of the variance in sales data compared to the other models. This indicates a higher level of accuracy and reliability in predictions.

The MAPE for the enhanced CNN-TCN model was the lowest among all models, at 2.0%. This shows that the CNN-TCN model was particularly effective at making accurate predictions relative to the actual sales values, making it the best performer in this aspect. The PICP was highest for the CNN-TCN model at 96%, indicating that this model provided the most reliable prediction intervals, which is crucial for making business decisions under uncertainty.

The introduction of ResNet and TCN layers not only improved accuracy but also enhanced the model's robustness. The use of residual connections in the ResNet architecture helped in mitigating the vanishing gradient problem, enabling the network to train effectively even with deep architectures. The dilated convolutions in the TCN model allowed for better handling of long-range dependencies, which is essential in time-series forecasting where past events can influence future outcomes over extended periods.

6. Conclusion

In this study, we have extensively explored the application of advanced hybrid deep learning architectures, specifically Convolutional Neural Networks (CNNs) integrated with Residual Networks (ResNets) and Temporal Convolutional Networks (TCNs), to enhance sales forecasting in dynamic market environments. By leveraging the "Store Item Demand Forecasting Challenge" dataset, which provides a rich and complex source of sales data across multiple stores and items over five years, we demonstrated the significant improvements in forecasting accuracy and reliability achieved by these enhanced architectures.

Accurate sales forecasting is critical for businesses, particularly in sectors like retail and supply chain management, where inventory control, resource allocation, and strategic planning hinge on the ability to predict future demand accurately. Traditional forecasting methods, though effective in certain contexts, often fall short in adapting to the complexities of modern markets characterized by rapid changes and non-linear consumer behavior patterns.

Our research makes a substantial contribution to the field by proposing and validating an advanced hybrid architecture that combines the strengths of CNNs, ResNets, and TCNs. This approach effectively addresses the limitations of standard CNN models in handling long-range dependencies and complex temporal patterns, which are prevalent challenges in time-series forecasting. The integration of ResNet's residual blocks mitigates the vanishing gradient problem, allowing for the training of deeper and more effective networks. Simultaneously, the TCN layers enhance the model's ability to capture long-term dependencies through dilated convolutions, making it particularly suitable for managing the sequential nature of sales data.

Empirical results from our experiments clearly demonstrate the superiority of the enhanced CNN-ResNet and CNN-TCN architectures over traditional CNN models and other deep learning approaches, such as LSTM. The enhanced models achieved lower Mean Absolute Error (MAE) and Root Mean Square Error (RMSE) values, indicating a significant reduction in both average and large forecasting errors. The R-squared values were closer to 1, reflecting a higher proportion of explained variance in the sales data, thus indicating greater prediction accuracy.

Furthermore, the CNN-TCN model demonstrated the lowest Mean Absolute Percentage Error (MAPE) at 2.0

Despite the promising results, there are several limitations that warrant discussion. First, the model's performance is highly dependent on the quality and representativeness of the training data. Although the "Store Item Demand Forecasting Challenge" dataset is comprehensive, significant changes in market conditions, consumer behavior, or external economic factors not captured in the dataset could impact the model's predictive accuracy.

Second, the computational complexity of the enhanced architectures, particularly with the integration of ResNet and TCN layers, requires considerable processing power and longer training times. This may limit the model's applicability in real-time forecasting scenarios or in environments with constrained computational resources.

Lastly, while the model's performance metrics were favorable, the architecture was specifically tuned for this dataset. Applying the model to other datasets or industries may require additional customization and optimization to achieve similar performance levels.

The findings from this study open several avenues for future research. One promising direction is the exploration of hybrid architectures that combine CNNs, ResNets, and TCNs with other advanced techniques such as attention mechanisms or Transformer models. These methods could further enhance the model's ability to capture complex temporal patterns and dependencies, especially in highly dynamic and multi-dimensional datasets.

Another potential research area is the integration of reinforcement learning techniques with deep learning models to create adaptive forecasting systems that can continuously learn and improve as new data becomes available. This approach could be particularly valuable in real-time forecasting scenarios where the model needs to rapidly respond to changing conditions.

Additionally, future studies could focus on applying the proposed architecture across different industries and datasets to assess its generalizability and robustness. Exploring the use of alternative data sources, such as social media sentiment or weather patterns, as additional inputs to the model could also provide deeper insights into the factors influencing sales and further improve forecasting accuracy.

In conclusion, this study demonstrates the significant potential of advanced hybrid deep learning architectures in enhancing sales forecasting accuracy. By addressing the key limitations of traditional models and proposing innovative solutions, our research contributes to the ongoing development of more powerful and reliable forecasting tools, providing businesses with valuable insights for strategic decision-making in an increasingly complex and competitive market environment.

Declaration of Interests

The authors declare no conflicts of interest, financial or otherwise, regarding the publication of this paper. This research was conducted as part of a PhD thesis and was carried out independently, without any external influence. The authors have no affiliations or involvement with any organizations that could affect the objectivity of the work reported.

REFERENCES

1. F. Messaoudi, M. Loukili, and M. El Ghazi, "Demand Prediction Using Sequential Deep Learning Model," in 2023 International Conference on Information Technology (ICIT), pp. 577-582, IEEE, August 2023.
2. D. Suman, "Deep Learning Methods for Demand Time Series Forecasting," 2024.
3. M. Loukili, F. Messaoudi, and M. E. Ghazi, "Enhancing Customer Retention through Deep Learning and Imbalanced Data Techniques," Iraqi Journal of Science, vol. 65, no. 5, pp. 2853-2866, 2024.
4. R. El Youbi, F. Messaoudi, and M. Loukili, "Machine Learning-driven Dynamic Pricing Strategies in E-Commerce," in 2023 14th International Conference on Information and Communication Systems (ICICS), pp. 1-5, IEEE, November 2023.
5. M. Loukili, F. Messaoudi, M. E. Ghazi, and H. Azirar, *Predicting Future Sales: A Machine Learning Algorithm Showdown*, In The International Conference on Artificial Intelligence and Smart Environment, pp. 26-31, Cham: Springer Nature Switzerland, 2023, November.
6. M. Loukili, and F. Messaoudi, "Enhancing Cold-Start Recommendations with Innovative Co-SVD: A Sparsity Reduction Approach," Statistics, Optimization & Information Computing, vol. 13, no. 1, pp. 396-408, 2025.
7. M. I. A. Efat, P. Hajek, M. Z. Abedin, R. U. Azad, M. A. Jaber, S. Aditya, and M. K. Hassan, "Deep-learning model using hybrid adaptive trend estimated series for modelling and forecasting sales," Annals of Operations Research, pp. 1-32, 2022.
8. F. Messaoudi, and M. Loukili, "E-commerce Personalized Recommendations: a Deep Neural Collaborative Filtering Approach," in Operations Research Forum, vol. 5, no. 1, pp. 1-25, Springer International Publishing, March 2024.
9. V. Shah, and S. Dimitrov, "A comparative study of univariate time-series methods for sales forecasting," International Journal of Business and Data Analytics, vol. 2, no. 2, pp. 187-216, 2022.
10. Y. Tang, J. Kurths, W. Lin, E. Ott, and L. Kocarev, "Introduction to focus issue: When machine learning meets complex systems: Networks, chaos, and nonlinear dynamics," Chaos: An Interdisciplinary Journal of Nonlinear Science, vol. 30, no. 6, 2020.
11. M. Loukili, and F. Messaoudi, "Machine Learning, Deep Neural Network and Natural Language Processing Based Recommendation System," in International Conference on Advanced Intelligent Systems for Sustainable Development, pp. 65-76, Springer Nature Switzerland, May 2022.
12. S. Mäkelä, "Optimizing the inventory system of a pharmaceutical company to improve efficiency-A Case Study," 2022.
13. M. Arunkumar, S. Palaniappan, R. Sujithra, and S. VijayPrakash, "Exploring Time Series Analysis Techniques for Sales Forecasting," in International Conference on Data Science and Network Engineering, pp. 41-55, Springer Nature Singapore, June 2023.
14. W. H. Chung, Y. Gu, and S. Yoo, "District Heater Load Forecasting Based on Machine Learning and Parallel CNN-LSTM Attention," SSRN Electronic Journal, 2022.
15. F. R. Adaryani, S. J. Mousavi, and F. Jafari, "Short-term Rainfall Forecasting using Machine Learning-Based Approaches of PSO-SVR, LSTM and CNN," Journal of Hydrology, 2022.
16. J. M. Ahn, J. Kim, and K. Kim, "Ensemble Machine Learning of Gradient Boosting (XGBoost, LightGBM, CatBoost) and Attention-Based CNN-LSTM for Harmful Algal Blooms Forecasting," Toxins, 2023.

17. S. Malakouti, A. Ghiasi, A. A. Ghavifekr, and P. Emami, ”Predicting wind power generation using machine learning and CNN-LSTM approaches,” Wind Engineering, 2022.
18. D. M. U. Ashraf, ”A Predictive Analysis of Retail Sales Forecasting using Machine Learning Techniques,” Journal of Computer Science and Information Technology, 2022.
19. L. Eglite, and I. Birzniece, ”Retail Sales Forecasting Using Deep Learning: Systematic Literature Review,” Complex Systems Informatics and Modeling Quarterly, 2022.
20. Y. Ensafi, S. H. Amin, G. Zhang, and B. Shah, ”Time-series forecasting of seasonal items sales using machine learning - A comparative analysis,” Journal of Industrial Information Integration, 2022.
21. R. Garg, S. Barpanda, S. G. SalankeN, and S. Ramya, ”Machine Learning Algorithms for Time Series Analysis and Forecasting,” arXiv preprint arXiv:2211.14387, 2022.
22. B. S. Suresh, and M. Suresh, ”A Comprehensive Analysis of Retail Sales Forecasting Using Machine Learning and Deep Learning Methods,” 2023 International Conference on Data Science and Network Security (ICDSNS), 2023.
23. Ş. T. Özçelik, F. B. Tek, and E. Şekerci, ”Hotel Sales Forecasting with LSTM and N-BEATS,” 2023 7th International Conference on Computer Science and Engineering (UBMK), 2023.