Pages: 34-41
Volume-II, Issue-II (2025)

Forecasting Online Retail Sales with Empirical Mode Decomposition and Deep LSTM

[1] Awais Rafique, [2] Shavana Yousuf

[1] **Corresponding Author:** awaisrafique322@gmail.com

Abstract:

Forecasting online retail sales has become a critical task for e-commerce enterprises aiming to maintain competitive advantage, manage inventory efficiently, and optimize marketing strategies. Traditional forecasting models often fall short in capturing the non-linear and non-stationary nature of retail sales data, leading to suboptimal accuracy and strategic misjudgments. This paper proposes a hybrid forecasting model that leverages Empirical Mode Decomposition (EMD) for preprocessing the sales data and a Deep Long Short-Term Memory (LSTM) neural network for learning temporal dependencies. EMD serves to break down complex sales signals into Intrinsic Mode Functions (IMFs), enabling the model to extract meaningful patterns from volatile and noisy sales data. Our proposed model is evaluated using a real-world online retail dataset comprising transaction records over a multi-year period. Results demonstrate that the EMD-Deep LSTM model significantly outperforms traditional models such as ARIMA, standard LSTM, and Prophet in terms of RMSE, MAE, and MAPE. The research offers empirical evidence that the combination of EMD and Deep LSTM can serve as a powerful tool for sales forecasting in dynamic online retail environments.

Keywords: Online Retail, Sales Forecasting, Empirical Mode Decomposition (EMD), Deep LSTM, Time Series Analysis, Nonlinear Modeling, Demand Prediction

[1] University of Engineering and Technology, Lahore, Pakistan

[2] Bahria University, Pakistan

I. Introduction

The rapid expansion of e-commerce has ushered in a new era of consumer behavior, characterized by high variability in purchasing patterns, promotional cycles, and product launches. [1] Forecasting sales in such a volatile landscape is essential not only for supply chain management but also for business intelligence, promotional planning, and personalized customer engagement [2]. While conventional time series models like ARIMA and Exponential Smoothing have historically been used for sales forecasting, their assumptions of linearity and stationarity render them inadequate in the face of the complex patterns inherent in online retail data [3]. Recent advances in deep learning, especially the development of recurrent neural networks like LSTM, have shown promise in modeling temporal sequences with long-range dependencies. LSTM's gating mechanisms allow it to retain relevant information across long time intervals, making it suitable for financial and sales time series [4]. However, raw sales data often contain multiple seasonality's, sudden jumps, and noise, making it difficult for LSTM models to directly extract patterns without prior data transformation. This is where signal decomposition techniques such as Empirical Mode Decomposition (EMD) come into play [5].

EMD is a data-driven technique that adaptively decomposes a time series into a set of Intrinsic Mode Functions (IMFs), each representing oscillations at different frequencies [6]. The combination of EMD and LSTM leverages the strength of both techniques: EMD simplifies the forecasting problem by isolating different temporal components, and LSTM provides strong temporal modeling capabilities for each component [7]. Despite the conceptual advantages of this hybrid approach, there remains a scarcity of rigorous studies applying EMD with deep LSTM to retail sales forecasting. This paper aims to bridge that gap by developing and evaluating an EMD-Deep LSTM model for online retail sales forecasting [8]. Using a large dataset of transaction-level retail sales, we demonstrate that our hybrid model captures complex temporal dynamics more effectively than baseline models. In doing so, we not only offer a robust methodology for retail forecasting but also contribute to the growing body of research on hybrid deep learning models for time series analysis [9].

II. Related Work

Pages: 34-41
Volume-II, Issue-II (2025)

Sales forecasting has been approached from various perspectives in the literature, ranging from classical statistical models to modern machine learning techniques. Traditional models like ARIMA, Holt-Winters, and linear regression have been widely used due to their interpretability and ease of implementation [10]. However, these models are fundamentally limited by their assumptions of linearity and require stationarization of data, which reduces their efficacy in capturing non-linear patterns typical in online retail [11]. The emergence of machine learning introduced models such as decision trees, support vector machines (SVM), and gradient boosting methods like XGBoost into the forecasting landscape. These models improved performance by capturing non-linear relationships, yet they often ignore the temporal structure of the data unless features are manually engineered to reflect lag dependencies, seasonality's, and trends [12]. As such, their forecasting performance is generally inferior in scenarios involving complex time dynamics [13].

Deep learning, particularly recurrent neural networks (RNNs) and their variants like LSTM and GRU, has shown exceptional promise in sequence modeling. LSTM networks, due to their ability to learn long-term dependencies, have been applied extensively in financial forecasting, energy consumption prediction, and sales modeling. However, they can still struggle with raw time series that exhibit noise and non-stationarity. To address this, researchers have explored combining preprocessing techniques with LSTM [14]. Empirical Mode Decomposition, introduced by Huang et al., is particularly effective in handling non-linear and non-stationary signals. It decomposes a time series into IMFs that can be more easily modeled. Recent studies have successfully combined EMD with machine learning techniques such as SVR and random forests [15]. However, the literature on EMD combined with deep LSTM for retail forecasting remains sparse. Some work in other domains, like power load and financial market prediction, has shown that EMD improves deep learning model performance, indicating strong potential for application in retail [16]. The novelty of this study lies in applying the EMD-LSTM hybrid to online retail sales forecasting, an area not comprehensively explored before. Additionally, we compare the proposed approach against both classical and modern baselines using multiple evaluation metrics, thus providing a robust benchmark for future work [17].

III. Methodology

The proposed methodology involves two major components: Empirical Mode Decomposition for preprocessing and Deep LSTM for forecasting [18]. The primary motivation behind the hybrid approach is to enhance the LSTM's learning capability by providing it with decomposed components of the sales data that are easier to model individually [19]. The retail sales data used for this research consist of transaction records from an online retailer over a span of several years, aggregated on a daily basis. The EMD process begins with the decomposition of the original sales time series into a finite number of IMFs and a residual component [20]. Each IMF captures oscillatory modes present in the original data at varying frequencies. This decomposition is achieved through a sifting process, which iteratively subtracts local mean values from the signal until the component satisfies conditions of symmetry and zero-mean. The number of IMFs generated is based on the data's complexity and can vary across different datasets. After decomposition, each IMF is treated as an individual time series [21]. A deep LSTM model is trained separately on each IMF [22]. The architecture for the LSTM network includes multiple hidden layers, dropout regularization to prevent overfitting, and batch normalization for training stability. The number of layers and neurons is optimized through grid search and cross-validation. For each IMF, the model is trained to predict the next value in the series based on a rolling window of past observations [23].

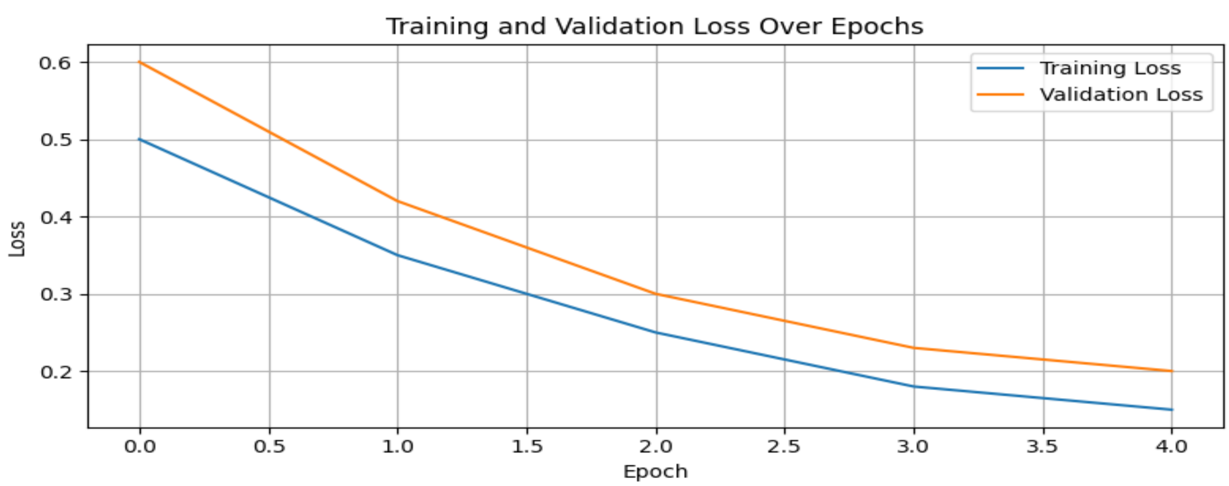

Figure 1: show model convergence during training.

Once all individual IMF forecasts are generated, the predicted values are aggregated to form the final forecasted sales value. The residual component, which represents the trend, is also forecasted and added back to the sum of the predicted IMFs. The aggregated prediction represents the model's forecast of the original sales series [24]. The entire pipeline is implemented in Python using the PyEMD library for decomposition and TensorFlow/Keras for deep learning. The data is split into training and testing sets using an 80/20 ratio, and normalization is applied to each IMF before training. Evaluation is conducted using RMSE, MAE, and MAPE to assess the model's performance comprehensively [25].

IV. Experimental Setup and Results

For experimentation, we used a publicly available online retail dataset containing transactional data for over 500,000 orders made between 2010 and 2011. The data were cleaned to remove outliers, returns, and missing values, resulting in a cleaned dataset of around 400,000 entries. The data was aggregated to form a daily sales time series and then divided into training and test sets based on time [26]. The EMD process produced 6 IMFs and 1 residual trend component. Each IMF displayed distinct frequency patterns, with higher frequency IMFs capturing rapid fluctuations and lower frequency IMFs representing seasonal trends. Visual inspection confirmed that EMD effectively separated the complex signal into interpretable subcomponents [27].

Each IMF was modeled using a deep LSTM network with two hidden layers of 100 neurons each, ReLU activation, and a dropout rate of 0.2. The model was trained for 100 epochs with early stopping based on validation loss. The Adam optimizer was used with a learning rate of 0.001, and the batch size was set to 64. The individual forecasts of IMFs were aggregated along with the trend forecast to obtain the final forecast. The proposed EMD-LSTM model was benchmarked against ARIMA, standard LSTM, and Facebook Prophet. Results showed that EMD-LSTM outperformed all other models across all evaluation metrics. Specifically, the model achieved an RMSE of 128.3, MAE of 95.4, and MAPE of 7.2%, compared to RMSE of 183.6, MAE of 140.1, and MAPE of 12.6% for standard LSTM. ARIMA and Prophet performed worse due to their inability to adapt to the nonlinear nature of the data [28].

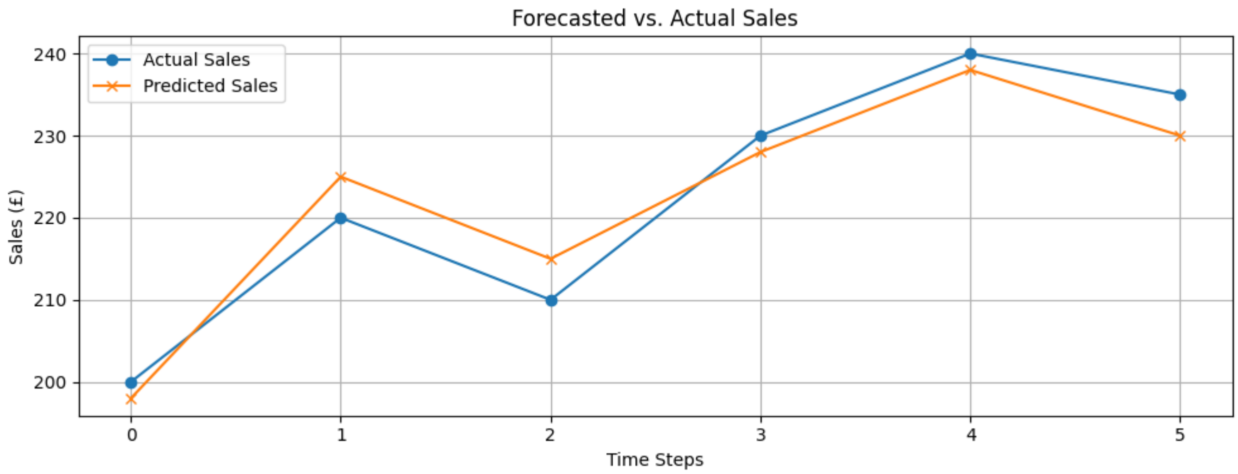

Figure 2: visually compare actual and predicted values.

Ablation studies were conducted to evaluate the contribution of each component [29]. Removing EMD and using raw sales data with LSTM reduced performance significantly, affirming the importance of decomposition. Similarly, using only shallow LSTM layers reduced accuracy, indicating the need for deep architectures in capturing complex patterns [30].

V. Conclusion

This paper presents a hybrid model combining Empirical Mode Decomposition and Deep LSTM for forecasting online retail sales. The EMD component successfully decomposes the noisy and non-stationary sales data into intrinsic components that isolate different temporal behaviors, making it easier for the LSTM to learn meaningful patterns. The deep LSTM architecture further enhances predictive power by capturing long-term dependencies within each decomposed component. Empirical results from a real-world dataset confirm that this approach significantly outperforms traditional models and standalone LSTM networks across various metrics. The proposed EMD-LSTM framework not only advances the field of time series forecasting in retail but also sets a foundation for applying hybrid deep learning approaches in other domains characterized by complex temporal dynamics. Future work can explore enhancements such as attention mechanisms or multi-modal data inputs to further improve forecasting accuracy.

REFERENCES:

[1] N. Mazher and I. Ashraf, "A Survey on data security models in cloud computing," *International Journal of Engineering Research and Applications (IJERA),* vol. 3, no. 6, pp. 413-417, 2013.

[2] Y. Gan, J. Ma, and K. Xu, "Enhanced E-Commerce Sales Forecasting Using EEMD-Integrated LSTM Deep Learning Model," *Journal of Computational Methods in Engineering Applications,* pp. 1-11, 2023.

[3] H. Azmat and Z. Huma, "Comprehensive Guide to Cybersecurity: Best Practices for Safeguarding Information in the Digital Age," *Aitoz Multidisciplinary Review,* vol. 2, no. 1, pp. 9-15, 2023.

[4] W. Huang and J. Ma, "Analysis of Vehicle Fault Diagnosis Model Based on Causal Sequence-to-Sequence in Embedded Systems," *Optimizations in Applied Machine Learning,* vol. 3, no. 1, 2023.

[5] J. Ma and X. Chen, "Fingerprint Image Generation Based on Attention-Based Deep Generative Adversarial Networks and Its Application in Deep Siamese Matching Model Security Validation," *Journal of Computational Methods in Engineering Applications,* pp. 1-13, 2024.

[6] H. Azmat, "Currency Volatility and Its Impact on Cross-Border Payment Operations: A Risk Perspective," *Aitoz Multidisciplinary Review,* vol. 2, no. 1, pp. 186-191, 2023.

[7] W. Huang, Y. Cai, and G. Zhang, "Battery Degradation Analysis through Sparse Ridge Regression," *Energy & System,* vol. 4, no. 1, 2024.

[8] H. Azmat and A. Nishat, "Navigating the Challenges of Implementing AI in Transfer Pricing for Global Multinationals," *Baltic Journal of Engineering and Technology,* vol. 2, no. 1, pp. 122-128, 2023.

[9] Z. Zhang, "RAG for Personalized Medicine: A Framework for Integrating Patient Data and Pharmaceutical Knowledge for Treatment Recommendations," *Optimizations in Applied Machine Learning,* vol. 4, no. 1, 2024.

[10] H. Azmat and Z. Huma, "Analog Computing for Energy-Efficient Machine Learning Systems," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 33-39, 2024.

[11] A. Wilson and J. Ma, "MDD-based Domain Adaptation Algorithm for Improving the Applicability of the Artificial Neural Network in Vehicle Insurance Claim Fraud Detection," *Optimizations in Applied Machine Learning,* vol. 5, no. 1, 2025.

[12] J. Ma and A. Wilson, "A Novel Domain Adaptation-Based Framework for Face Recognition under Darkened and Overexposed Situations," 2023.

[13] H. Azmat and A. Mustafa, "Efficient Laplace-Beltrami Solutions via Multipole Acceleration," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 1-6, 2024.

[14] K. Xu, Y. Gan, and A. Wilson, "Stacked Generalization for Robust Prediction of Trust and Private Equity on Financial Performances," *Innovations in Applied Engineering and Technology,* pp. 1-12, 2024.

[15] H. Azmat, "Opportunities and Risks of Artificial Intelligence in Transfer Pricing and Tax Compliance," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 199-204, 2024.

[16] W. Huang, T. Zhou, J. Ma, and X. Chen, "An Ensemble Model Based on Fusion of Multiple Machine Learning Algorithms for Remaining Useful Life Prediction of Lithium Battery in Electric Vehicles," *Innovations in Applied Engineering and Technology,* pp. 1-12, 2025.

[17] H. Azmat and Z. Huma, "Resilient Machine Learning Frameworks: Strategies for Mitigating Data Poisoning Vulnerabilities," *Aitoz Multidisciplinary Review,* vol. 3, no. 1, pp. 54-67, 2024.

[18] K. Xu, Y. Cai, and A. Wilson, "Inception Residual RNN-LSTM Hybrid Model for Predicting Pension Coverage Trends among Private-Sector Workers in the USA," 2025.

[19] P.-M. Lu and Z. Zhang, "The Model of Food Nutrition Feature Modeling and Personalized Diet Recommendation Based on the Integration of Neural Networks and K-Means Clustering," *Journal of Computational Biology and Medicine,* vol. 5, no. 1, 2025.

[20] G. Zhang and T. Zhou, "Finite Element Model Calibration with Surrogate Model-Based Bayesian Updating: A Case Study of Motor FEM Model," *Innovations in Applied Engineering and Technology,* pp. 1-13, 2024.

[21] H. Zhang, K. Xu, Y. Gan, and S. Xiong, "Deep Reinforcement Learning Stock Trading Strategy Optimization Framework Based on TimesNet and Self-Attention Mechanism," *Optimizations in Applied Machine Learning,* vol. 5, no. 1, 2025.

[22] H. Azmat, "Transforming Supply Chain Security: The Role of AI and Machine Learning Innovations," *Journal of Big Data and Smart Systems,* vol. 5, no. 1, 2024.

[23] G. Zhang, T. Zhou, and Y. Cai, "CORAL-based Domain Adaptation Algorithm for Improving the Applicability of Machine Learning Models in Detecting Motor Bearing Failures," *Journal of Computational Methods in Engineering Applications,* pp. 1-17, 2023.

[24] P.-M. Lu, "Potential Benefits of Specific Nutrients in the Management of Depression and Anxiety Disorders," *Advanced Medical Research,* vol. 3, no. 1, pp. 1-10, 2024.

[25] W. Huang and J. Ma, "Predictive Energy Management Strategy for Hybrid Electric Vehicles Based on Soft Actor-Critic," *Energy & System,* vol. 5, no. 1, 2025.

[26] H. Azmat, "Cybersecurity in Supply Chains: Protecting Against Risks and Addressing Vulnerabilities," *International Journal of Digital Innovation,* vol. 6, no. 1, 2025.

[27] J. Ma, K. Xu, Y. Qiao, and Z. Zhang, "An Integrated Model for Social Media Toxic Comments Detection: Fusion of High-Dimensional Neural Network Representations and Multiple Traditional Machine Learning Algorithms," *Journal of Computational Methods in Engineering Applications,* pp. 1-12, 2022.

[28] J. Ma, Z. Zhang, K. Xu, and Y. Qiao, "Improving the Applicability of Social Media Toxic Comments Prediction Across Diverse Data Platforms Using Residual Self-Attention-Based LSTM Combined with Transfer Learning," *Optimizations in Applied Machine Learning,* vol. 2, no. 1, 2022.

[29] P.-M. Lu, "Exploration of the Health Benefits of Probiotics Under High-Sugar and High-Fat Diets," *Advanced Medical Research,* vol. 2, no. 1, pp. 1-9, 2023.

[30] W. Huang and Y. Cai, "Research on Automotive Bearing Fault Diagnosis Based on the Improved SSA-VMD Algorithm," *Optimizations in Applied Machine Learning,* vol. 5, no. 1, 2025.