2026 Volume: 02 Issue: 05 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-07

TIME SERIES-DRIVEN E-COMMERCE SALES FORECASTING AND INTELLIGENT RECOMMENDATION SYSTEM

*[1]Bhavani S., [1]Nanditha B.,[1]Deepthi V. T., [2]Kaveri A.

[1]Dept of CSE, Jyothy Institute of Technology, Bangalore.

[2]Assistant Professor, Jyothy Institute of Technology, Bangalore.

Article Received: 26 March 2026

Article Revised: 16 April 2026

Published on: 06 May 2026

***Corresponding Author: Bhavani S.**

Dept of CSE, Jyothy Institute of Technology, Bangalore.

DOI: https://doi-doi.org/101555/ijrpa.7810

ABSTRACT

E-commerce platforms generate large volumes of data, making accurate sales forecasting essential for effective decision-making. Traditional models such as ARIMA can capture trends and seasonality but struggle with complex patterns in real-world data. Recent advancements in machine learning and deep learning, including models like LSTM and XGBoost, have significantly improved forecasting accuracy. This paper presents a survey of various forecasting approaches, including statistical methods, machine learning techniques, and hybrid models that combine their strengths. The study highlights the advantages and limitations of these methods and emphasizes the growing importance of hybrid and AI-driven approaches. It also identifies research gaps and discusses future directions for developing more efficient and scalable forecasting systems in e-commerce.

KEYWORDS: E-commerce, Time Series Forecasting, LSTM, XGBoost, Machine Learning, Deep Learning, Hybrid Models

INTRODUCTION

Ecommerce has rapidly evolved into one of the most significant sectors in the digital economy, generating vast amounts of data related to sales, customer behavior, and product demand. Analyzing this data effectively is essential for businesses to make informed decisions, optimize inventory, and improve customer satisfaction. One of the major challenges in e-commerce is accurately predicting future sales, as it is influenced by various factors such as trends, seasonality, pricing strategies, and changing customer preferences [1], [6].

Traditional time series forecasting methods such as ARIMA and SARIMA have been widely used due to their ability to model trends and seasonal patterns [4], [9]. However, these models are limited in handling complex nonlinear relationships and large- scale data commonly found in modern e-commerce systems. With the advancement of artificial intelligence, machine learning and deep learning techniques have emerged as powerful alternatives that can capture hidden patterns and improve prediction accuracy [5], [7].

Models such as Long Short-Term Memory (LSTM), Gated Recurrent Units (GRU), and XGBoost have shown significant improvements in forecasting performance by effectively learning temporal dependencies and feature relationships. In recent years, hybrid approaches that combine statistical models with machine learning techniques have gained attention, as they leverage the strengths of both methods to achieve better results.

This paper focuses on analyzing various forecasting techniques used in e-commerce and highlights the transition from traditional models to advanced AI- based approaches. It also emphasizes the importance of developing efficient and scalable systems that can handle real-world complexities and support better decision-making in dynamic business environments.

LITERATURE SURVEY

Recent research in e-commerce sales forecasting has evolved from traditional statistical approaches to advanced machine learning and deep learning techniques. Earlier studies mainly relied on time series models such as ARIMA and SARIMA, which are effective in capturing linear trends and seasonal patterns. These models are simple, interpretable, and require less computational effort, but they are limited in handling complex nonlinear relationships and large-scale datasets.

With the advancement of artificial intelligence, machine learning models such as XGBoost, Random Forest, and Linear Regression have been widely used for forecasting tasks. These models can analyze multiple influencing factors such as pricing, promotions, and external conditions, leading to improved prediction accuracy. However, they do not inherently capture temporal dependencies, which are essential in time series data.

To overcome these limitations, deep learning models such as Long Short-Term Memory (LSTM) and Gated Recurrent Units (GRU) have been introduced. These models are specifically designed to process sequential data and are capable of capturing long-term dependencies and nonlinear patterns. As a result, they have shown superior performance in complex forecasting scenarios. However, they require large datasets, higher computational resources, and careful tuning.

Recent studies focus on hybrid approaches that combine statistical models with machine learning or deep learning techniques. For example, combining ARIMA with LSTM helps capture both linear and nonlinear patterns, while integrating XGBoost with LSTM allows simultaneous learning of feature relationships and temporal dependencies. These hybrid models have demonstrated improved accuracy and robustness compared to individual models. In addition, some research incorporates external factors such as holidays, weather conditions, and economic indicators to further enhance prediction performance. Real-time forecasting systems have also been explored, enabling dynamic predictions based on live data streams.

Despite these advancements, challenges such as model complexity, data quality, and lack of integration with intelligent decision-making systems still exist. Overall, the literature shows a clear transition toward hybrid and AI-based approaches for achieving more accurate and scalable forecasting solutions.

COMPARATIVE ANALYSIS

Comparative analysis of existing forecasting techniques reveals that each approach has its own strengths and limitations depending on the nature of the data and application requirements. Traditional statistical models such as ARIMA and SARIMA are widely used due to their simplicity and ability to model linear trends and seasonal patterns. These models are easy to interpret and require less computational effort, making them suitable for small and well-structured datasets. However, they are not effective in handling complex nonlinear relationships and high-dimensional data, which are common in modern e-commerce systems.

Machine learning models such as XGBoost, Random Forest, and Linear Regression have been introduced to overcome some of these limitations.

These models can analyze multiple input features simultaneously, including pricing, promotions, and external factors, leading to improved forecasting performance. They are particularly effective in capturing feature interactions and handling structured data. However, a major drawback is that they do not inherently consider the sequential nature of time series data, which can reduce their effectiveness in capturing temporal dependencies. Deep learning models such as LSTM and GRU are specifically designed to process sequential data and have shown strong performance in time series forecasting. These models can capture long-term dependencies, nonlinear patterns, and complex relationships within the data. As a result, they provide higher accuracy in dynamic and large-scale datasets. However, they require large amounts of training data, high computational resources, and careful parameter tuning, which can make them difficult to implement and optimize.

Hybrid models have emerged as the most effective solution by combining the strengths of multiple approaches. For instance, integrating ARIMA with LSTM allows the system to capture both linear and nonlinear patterns, while combining XGBoost with LSTM enables simultaneous learning of feature relationships and temporal dependencies. These models provide better accuracy, robustness, and adaptability compared to standalone methods. However, they increase system complexity and require more effort in model design, training, and maintenance.

Additionally, recent approaches incorporate external factors such as weather conditions, holidays, and economic indicators, which significantly improve forecasting performance. Real-time forecasting systems are also gaining attention, as they allow continuous updates and dynamic predictions based on live data streams. Despite these advancements, challenges such as data quality, scalability, and computational cost remain significant.

Overall, the comparison indicates that while traditional and individual models have their advantages, hybrid and AI-driven approaches are more suitable for real-world e-commerce forecasting due to their ability to handle complexity, improve accuracy, and adapt to changing data patterns.

RESEARCH GAP

Existing research in e-commerce sales forecasting has made significant progress using statistical, machine learning, and deep learning models. However, many traditional models such as ARIMA and SARIMA fail to effectively handle nonlinear patterns and large-scale dynamic data. While machine learning models improve prediction accuracy by analyzing multiple features, they often ignore temporal dependencies, which limits their effectiveness in time series forecasting.

Deep learning models such as LSTM and GRU address temporal relationships but require large datasets and high computational resources, making them less practical for real-time and resource- constrained environments. Although hybrid models attempt to combine the strengths of different approaches, they often increase system complexity and are not optimized for scalability, ease of implementation, or deployment in real-world systems.

Furthermore, most existing research focuses primarily on improving prediction accuracy, but fails to provide integrated solutions that combine forecasting with intelligent decision-support mechanisms. Many models do not effectively incorporate real-time data streams, and there is limited emphasis on adapting to rapidly changing market conditions. Additionally, the use of external influencing factors is often not unified within a single framework, reducing the overall effectiveness of predictions.

Another important limitation is the lack of user- oriented design and practical deployment strategies, which makes many proposed models difficult to implement in real business environments. Therefore, there is a need for a more efficient, scalable, and integrated system that not only improves forecasting accuracy but also supports real-time adaptability, intelligent decision-making, and practical usability in e-commerce applications.

PROPOSED IDEA

The proposed system aims to develop an efficient and intelligent framework for e-commerce sales forecasting by integrating time series analysis with advanced machine learning and deep learning techniques. The system utilizes historical sales data to identify important patterns such as trends, seasonality, and demand fluctuations. Models like LSTM are employed to capture temporal dependencies in sequential data, while XGBoost is used to analyze feature relationships and improve overall prediction accuracy. By combining these approaches, the system is able to effectively handle both linear and nonlinear patterns present in real- world datasets.

In addition to improving prediction accuracy, the system focuses on adaptability by incorporating real-time data processing. This allows the model to update forecasts dynamically as new data becomes available, making it suitable for rapidly changing e-commerce environments. The system also considers external factors such as promotions, seasonal effects, and customer behavior to enhance the reliability of predictions.

Furthermore, the proposed approach supports better decision-making by generating meaningful insights from the predicted data. These insights can be used for inventory management, demand planning, and business strategy optimization. Overall, the system aims to provide a scalable, accurate, and practical solution that overcomes the limitations of existing standalone models and can be effectively applied in real-world e-commerce applications.

CONCLUSION

This paper presented a comprehensive survey of various techniques used in e-commerce sales forecasting, including traditional statistical models, machine learning approaches, deep learning methods, and hybrid frameworks. The study highlighted the evolution of forecasting techniques from simple models such as ARIMA and SARIMA, which are effective in capturing linear trends and seasonality, to more advanced approaches capable of handling complex and dynamic data patterns. While traditional models are easy to interpret and implement, they are limited in their ability to process nonlinear relationships and large-scale datasets.

With the advancement of artificial intelligence, machine learning and deep learning models such as XGBoost, LSTM, and GRU have significantly improved forecasting performance. These models are capable of capturing hidden patterns, feature interactions, and temporal dependencies, making them more suitable for real-world e-commerce environments. However, they also introduce challenges such as higher computational requirements, the need for large datasets, and increased model complexity.

The comparative analysis emphasized that hybrid approaches, which combine statistical methods with machine learning and deep learning techniques, provide better accuracy and robustness compared to standalone models. By leveraging the strengths of multiple approaches, hybrid models can effectively handle both linear and nonlinear patterns while improving prediction reliability. Despite these improvements, issues such as scalability, real-time adaptability, and practical deployment remain key challenges in existing systems.

The research gap identified in this study highlights the lack of integrated systems that combine accurate forecasting with intelligent decision-support mechanisms. Many existing models focus only on prediction accuracy without addressing real-time data processing, user interaction, and practical usability in business environments.

To address these limitations, the proposed system integrates time series analysis with advanced learning techniques to provide accurate, scalable, and adaptive forecasting. It also supports better decision-making by generating meaningful insights for inventory management, demand planning, and business optimization.

In conclusion, the study demonstrates that hybrid and AI-driven approaches are more effective for modern e-commerce forecasting applications. Future work can focus on developing more efficient, real-time, and user-friendly systems that can be easily deployed in real-world scenarios while maintaining high accuracy and performance.

REFERENCES

1. C. Wang and J. Wang, "Research on E- Commerce Inventory Sales Forecasting Model Based on ARIMA and LSTM Algorithm," *Mathematics*, vol. 13, no. 11, pp. 1–18, Nov. 2025.
2. I. Boukrouh, S. Idriri, F. Tayalati, A. Azmani, and L. Bouhassien, "A Hybrid Approach

for Sales Forecasting: Combining Deep Learning and Time Series Analysis," *International Journal of Engineering (IJE), Transactions A: Basics*, vol. 38, no. 4, pp. 859–870, Apr. 2025.

3. Z. Mohammed, C. Anas, and M. El Hammoumi, "A Hybrid Learning Framework for Forecasting Uncertainty and Adaptive Inventory Planning in Retail Supply Chains," *Supply Chain Analytics*, vol. 13, pp. 100180, 2026.
4. T. Zhou, "Improved Sales Forecasting Using Trend and Seasonality Decomposition with LightGBM," *arXiv preprint arXiv:2305.17201*, 2023.
5. V. I. Kontopoulou, A. D. Panagopoulos, I. Kakkos, and G. K. Matsopoulos, "A Review of ARIMA vs. Machine Learning Approaches for Time Series Forecasting in Data Driven Networks," *Future Internet*, vol. 15, no. 8, p. 255, Jul. 2023.
6. S. B. Kwarteng and P. A. Andreevich, "Comparative Analysis of ARIMA, SARIMA and Prophet Model in Forecasting," *Research & Development*, vol. 5, no. 4, pp. 110–120, Oct. 2024.
7. K. Douaioui, R. Oucheikh, O. Benmoussa, and C. Mabrouki, "Machine Learning and Deep Learning Models for Demand Forecasting in Supply Chain Management: A Critical Review," *Applied System Innovation*, vol. 7, no. 5, p. 93, Sep. 2024.
8. Y. Chen, X. Xie, Z. Pei, W. Yi, C. Wang, W. Zhang, and Z. Ji, "Development of a Time Series E- Commerce Sales Prediction Method for Short- Shelf-Life Products Using GRU- LightGBM," *Applied Sciences*, vol. 14, no. 2, p. 866, Jan. 2024.
9. H. Jatte, S. Belattar, and E. K. Haimoudi, "Optimization of Forecasting Performance in the Retail Sector Using Artificial Intelligence," *International Journal of Electrical and Computer Engineering*, vol. 14, no. 1, pp. 123–132, Feb. 2024.
10. P. Yadav and C. R. Barde, "Retail Real-Time Sales Prediction System Using LSTM and XGBoost," *International Journal of Innovative Research in Computer Science & Technology*, vol. 11, no. 3, pp. 45–52, May 2023.
11. B. Wang and A. B. M. Zain, "A Hybrid XGBoost-LSTM Framework for Supply Chain Demand Forecasting: Empirical Evidence from Retail Multi-Store Data," Journal of Cultural Analysis and Social Change, vol. 10, no. 4, pp. 4056–4073, Dec. 2025.
12. S. Mansur, K. Sattar, S. E. Hosseini, S. Pervez, Ahmad, K. Saleem, and A. Z. Elhendi, "Sales Forecasting for Retail Stores Using Hybrid Neural Networks and Sales-Affecting Variables," PeerJ Computer Science, vol. 11, p. e3058, Sep. 2025.