GEORGE BROWN PRESS
JOURNALS, BOOKS & PROCEEDINGS

Economics and Management Innovation

Article

Prediction Framework for E-Commerce Platform Sales Data Based on Informer: A Study on Furniture Sales on Amazon E-Commerce Platform

Lianzhen Wang [1,*], Xi Zhang [2] and Lu Jiang [3]

[1] Beijing City University, Beijing, 101399, China
[2] Booth School of Business, University of Chicago, Chicago, IL, USA
[3] Morgan Stanley, Fordham University, New York, NY, USA
* Correspondence: Lianzhen Wang, Beijing City University, Beijing, 101399, China

Abstract: Accurate sales forecasting is a critical requirement for e-commerce platforms, as it supports efficient inventory management, dynamic pricing, and targeted marketing strategies. Nonetheless, forecasting is inherently challenging due to the volatility of consumer demand, the impact of external factors such as economic indicators and weather conditions, and the occurrence of promotional events. To address these complexities, this study proposes a prediction framework based on the Informer model, a Transformer-based architecture specifically designed for long-sequence time series forecasting. The framework leverages weekly sales data of five furniture categories from the Amazon e-commerce platform spanning 2015 to 2020, combined with auxiliary features including price, discount rate, consumer price index (CPI), and weather-related variables. By integrating both temporal dependencies and exogenous influences, the model captures intricate patterns in sales dynamics. Experimental results demonstrate that the Informer model significantly outperforms traditional baselines, including LSTM, GRU, and Prophet, achieving a mean absolute error (MAE) of 37.3 and a root mean square error (RMSE) of 52.7. These findings highlight the model's effectiveness in enhancing e-commerce sales forecasting and underscore the potential of advanced deep learning techniques to improve operational decision-making for digital retail platforms.

Keywords: e-commerce; sales prediction; Informer; Amazon; time series forecasting; deep learning

Received: 09 October 2025
Revised: 23 October 2025
Accepted: 20 November 2025
Published: 26 November 2025

1. Introduction

In recent years, the rapid expansion of e-commerce platforms has profoundly reshaped consumer behavior and online retail management. Accurate sales forecasting has become a critical component of e-commerce operations, directly influencing inventory optimization, marketing strategy formulation, and revenue management. The ability to anticipate future demand allows platform operators to make data-driven decisions, minimize stockouts or overstock situations, and enhance overall operational efficiency. Nevertheless, sales prediction in the e-commerce domain remains a complex task due to the inherently dynamic nature of consumer demand and the multitude of external factors, including economic trends, seasonal variations, and promotional campaigns [1].

Traditional time series forecasting models, such as ARIMA, LSTM, GRU, and Prophet, have provided valuable insights into sales patterns; however, they often struggle to capture long-term temporal dependencies and the non-linear interactions among multiple influencing variables [2]. These limitations are especially pronounced in high-frequency, high-dimensional e-commerce datasets, which are characterized by irregular promotional events, volatile pricing, and rapidly changing consumer preferences. To overcome these challenges, this study proposes a prediction framework for e-commerce sales based on the Informer model, a Transformer-based architecture specifically designed for long-sequence time series forecasting. Unlike conventional recurrent models, the Informer utilizes a self-attention mechanism combined with a ProbSparse strategy, substantially enhancing computational efficiency and the capacity to process extensive input sequences.

For this study, weekly sales data of five furniture categories on the Amazon e-commerce platform covering 2015 to 2020 were collected and integrated with multi-source auxiliary features, including product price, discount rate, consumer price index (CPI), and weather indicators. This integration allows the model to capture both intrinsic temporal dynamics and exogenous environmental factors that influence online consumer behavior. Experimental results indicate that the proposed framework achieves superior forecasting performance, reducing the mean absolute error (MAE) to 37.3 and the root mean square error (RMSE) to 52.7, thereby significantly outperforming baseline models such as LSTM, GRU, and Prophet.

The main contributions of this study are threefold:

(1) A novel multi-source data integration framework that combines sales, economic, and environmental variables to enhance the contextual understanding of consumer demand in e-commerce.

(2) An optimized Informer-based prediction model that effectively captures long-term dependencies and improves accuracy in complex, volatile sales environments.

(3) Comprehensive empirical validation using real-world Amazon furniture sales data, providing actionable insights for inventory management, pricing strategies, and operational decision-making on digital retail platforms.

2. Literature Review

Sales forecasting on e-commerce platforms holds critical practical value, particularly for inventory management, pricing optimization, and marketing decision-making. On large-scale platforms such as Amazon, the sales of furniture products exhibit pronounced seasonality and volatility, directly impacting platform operational efficiency and the profit distribution among merchants. However, the inherently complex and non-linear characteristics of sales volume time-series data pose substantial challenges to both predictive accuracy and model generalization. In recent years, deep learning-based forecasting models have rapidly advanced, demonstrating clear advantages over traditional methods in handling high-dimensional, noisy, and irregular e-commerce datasets, and achieving significantly improved predictive performance.

Recent studies have explored various deep learning frameworks to enhance sales forecasting accuracy. For instance, some research leveraging the M5 dataset demonstrates that incorporating temporal features, event information, and product identifiers into a global DeepAR model consistently improves retail sales predictions. This approach achieves a 1.8% reduction in RMSSE and a 6.5% improvement in MASE compared to baseline models, clearly outperforming seasonal naïve benchmarks and providing a robust covariate-selection framework for subsequent deep learning forecasting studies [3]. Other approaches, such as the ATES-DNN hybrid model, first remove seasonality and holiday effects before allowing a deep neural network to learn spatio-temporal residual patterns. This methodology yields robust, noise-tolerant forecasts at the product-store level and outperforms state-of-the-art methods across forecasting horizons ranging from one to twelve months [4].

For short-life-cycle and new-product forecasting, strategies that enrich limited historical data via time-series clustering and data augmentation have proven effective. Benchmarking traditional statistical models such as ARIMAX against deep learning approaches including LSTM, GRU, and CNN on large datasets reveals that while ARIMAX can achieve lower MAE under ideal conditions (21–24% improvement), it is highly sensitive to added noise, whereas deep neural networks maintain robustness, offering clear guidance for model selection in practice [5]. Additionally, hybrid methods that combine signal decomposition techniques with deep learning, such as EEMD-LSTM, have been developed to tackle complex e-commerce sales patterns. By decomposing the original time series into intrinsic mode functions (IMFs) before applying deep prediction, these methods achieve high forecasting accuracy—up to 91% in the case of Amazon apparel data—while providing a benchmark for de-noising strategies, highlighting effective directions for further innovation in hybrid forecasting frameworks [6].

Overall, these studies illustrate the growing potential of deep learning and hybrid approaches for accurately modeling highly dynamic, non-linear, and high-dimensional e-commerce sales data. They underscore the importance of integrating temporal, event-driven, and exogenous features, as well as leveraging decomposition and noise-handling strategies, to enhance predictive performance in real-world online retail environments.

3. Data Introduction

The dataset used in this study comprises weekly sales records for five categories of furniture—chairs, tables, sofas, beds, and cabinets—on the Amazon e-commerce platform from 2015 to 2020. Each record contains detailed information including sales volume, product price, discount rate, economic indicators such as the consumer price index (CPI), and weather-related variables, such as average temperature and general weather conditions. The inclusion of both intrinsic product features and external environmental factors provides a comprehensive view of the determinants influencing consumer purchasing behavior, enabling more accurate modeling of sales dynamics.

To ensure the reliability and usability of the dataset, a series of preprocessing steps were conducted. Initially, missing values in the dataset, denoted as "-99" or "NaN," were systematically identified and removed to prevent potential distortion during model training. Categorical variables, such as weather conditions, were encoded into numerical representations suitable for deep learning models, while continuous variables, including price, discount rate, and CPI, were normalized to a consistent scale. Additionally, outliers were examined and handled through a combination of threshold-based filtering and statistical transformations to reduce the influence of anomalous records on the forecasting results. These preprocessing operations not only improved data quality but also facilitated more effective feature extraction and model convergence, ensuring that the subsequent forecasting models could learn robust and meaningful patterns from the data.

As shown in Figure 1, Amazon furniture sales exhibit clear seasonal patterns across the four quarters from 2015 to 2020. Sales peak during Fall and Winter, accounting for 26.04% of annual sales in each season, while Spring sales are relatively lower at 23.36%. Although the seasonal variation remains below 3%, a consistent "high in Fall/Winter, low in Spring/Summer" pattern emerges, highlighting the importance of incorporating seasonal components into time series forecasting models. Beyond seasonality, subtle trends over multiple years, such as gradual increases in overall sales volume and variations across product categories, are also observable, emphasizing the necessity for models that can capture both long-term trends and short-term fluctuations.

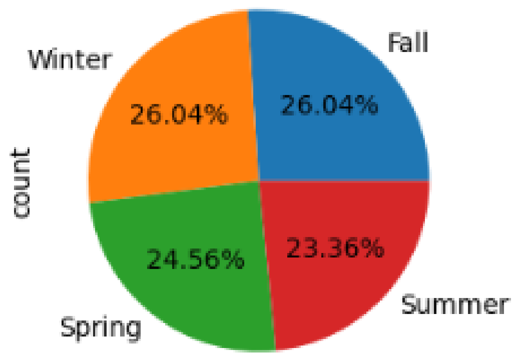

Figure 1. Sales of furniture in each of the four quarters from 2015 to 2020.

Overall, the processed dataset provides a robust foundation for conducting accurate and reliable time series forecasting experiments. By integrating multiple sources of information, addressing missing values and outliers, and standardizing features, the dataset ensures that predictive models are both effective and generalizable. Furthermore, the richness of the dataset allows for exploration of complex interactions among pricing, promotions, economic conditions, and weather factors, offering valuable insights for strategic decision-making in e-commerce operations, including inventory planning, dynamic pricing, and promotional campaign optimization.

4. Methodology

In this study, we designed an informant framework based on sparse attention mechanism for predicting weekly sales data of furniture on the Amazon e-commerce platform.

4.1. Feature Processing

In this study, we propose a sales forecasting framework for e-commerce furniture data based on the Informer deep learning architecture. Traditional recurrent models such as RNNs, LSTM, and GRU have demonstrated advantages in capturing short-term dependencies in time series data. However, when handling long sequences with multiple exogenous variables, these models often suffer from gradient vanishing and computational inefficiency. To overcome these limitations, we employ the Informer model, which is built on the Transformer framework and incorporates optimized attention mechanisms, making it particularly suitable for long sequence forecasting tasks in e-commerce scenarios.

At the input feature processing stage, we combine Amazon's weekly furniture sales data with external variables such as price, discount rate, consumer price index (CPI), and weather conditions. Let the raw time series be defined as:

$$\mathrm{X} = \{\mathrm{x}_1, \mathrm{x}_2, \dots, \mathrm{x}_\mathrm{T}\} \tag{1}$$

where $x_t \in R^d$ represent a d-dimensional feature vector at time step t, including sales volume, price, discount, economic indicators, and meteorological features. To eliminate scale differences, we apply normalization to all numerical variables:

$$\tilde{x}_t = \frac{x_t - min(x)}{max(x) - min(x)} \tag{2}$$

The preprocessed sequences are then embedded into high-dimensional vectors as inputs to the Informer.

4.2. ProbSparse Self-Attention

The Informer architecture introduces the ProbSparse Self-Attention mechanism to improve the efficiency of long sequence modeling [7]. Standard self-attention requires calculating similarity scores for all query-key pairs with a complexity of $O(L^2)$, where L is the sequence length. In contrast, ProbSparse Attention reduces the computational burden by retaining only the most informative attention scores, reducing complexity to $O(L\,log\,L)$). Formally, the self-attention mechanism can be expressed as:

$$Attention(Q,K,V) = Softmax(\frac{QK^T}{\sqrt{d_K}}) \tag{3}$$

while ProbSparse Attention applies a Top-u selection strategy to preserve only the largest attention values, thereby ensuring efficiency without sacrificing accuracy.

4.3. distilling operation

To further reduce redundancy in long input sequences, Informer incorporates a distilling operation in the encoder, which compresses sequence length using one-dimensional convolution and down-sampling [8]. This mechanism retains essential temporal dependencies while decreasing computational costs, making the model more adaptable to multi-feature, long-span e-commerce data. In the decoder, Informer adopts a generative decoding mechanism that can directly output predictions for multiple future steps. Given a forecasting horizon of length *h*, the model output is expressed as:

$$\hat{Y} = \{\hat{y}_{T+1}, \hat{y}_{T+1}, \dots, \hat{y}_{T+h}\} \tag{4}$$

where $\hat{y}_t \in R$ denotes the predicted furniture sales volume at week *t*.

In practical implementation, the multi-dimensional input features are first mapped into a unified vector space through embedding layers, combined with positional encoding to retain temporal information, and then fed into the encoder–decoder structure. The encoder extracts global dependencies between historical sales and exogenous features, while the decoder utilizes the attention mechanism to generate future sales sequences. Model parameters are optimized by minimizing the mean squared error (MSE) loss:

$$L = \frac{1}{h}\sum_{i-1}^{h}(y_{T+i} - \hat{y}_{T+i})^2, \tag{5}$$

where y_{T+i} is the ground-truth sales volume and $\hat{y}_{T+i}$ is the predicted value.

Through this design, the Informer model effectively captures the periodic and trend patterns of furniture sales while simultaneously incorporating the impact of external factors such as price, discount, economic indicators, and weather conditions. This enables superior accuracy and generalization ability in forecasting weekly furniture sales on the Amazon e-commerce platform.

4.4. Overall framework

Figure 2 illustrates the overall framework of the Informer model for e-commerce sales prediction. The model first integrates multiple input features, including historical sales, product price, discount rate, consumer price index (CPI), and weather information, which jointly characterize the dynamic patterns influencing furniture sales on the Amazon platform. These features are embedded and then fed into the encoder module, where temporal dependencies and correlations across multiple variables are effectively captured through the ProbSparse self-attention mechanism. The compressed and distilled representations are subsequently passed into the decoder, which generates multi-step predictions of future sales. This encoder–decoder architecture enables the Informer to model both long-term trends and short-term fluctuations, thereby providing accurate forecasts of weekly furniture sales in the e-commerce environment.

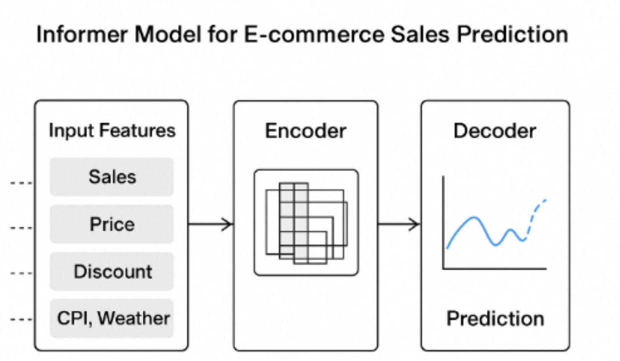

Figure 2. Model Structure.

5. Experiment

5.1. Evaluation Metrics

To comprehensively evaluate the performance of the proposed sales forecasting framework, several widely used error-based metrics for time series forecasting were adopted. Specifically, we used Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE).

These three metrics together provide a balanced assessment: MAE reflects the overall prediction deviation, RMSE penalizes large errors, and MAPE measures accuracy in percentage terms, which is particularly important for practical applications in e-commerce sales forecasting.

5.2. Results

Figure 3 shows the predictive effect of the Informer model on the weekly sales of furniture on the Amazon platform. It can be seen that during the training set stage (blue curve), the overall sales data shows some fluctuations, accompanied by a small number of peak outliers. In the prediction stage (where the green curve represents the predicted value and the orange curve represents the actual value), the model can track the trend of actual sales well. Especially in most stable intervals, the predicted values basically coincide with the true values, indicating that the model performs more accurately in capturing the long-term trend and short-term fluctuations of furniture sales. Although there may be some deviation during extreme peaks and sudden drops, the overall error margin is small, and the predicted curve can closely fit the actual curve, reflecting the effectiveness and robustness of the model in this task.

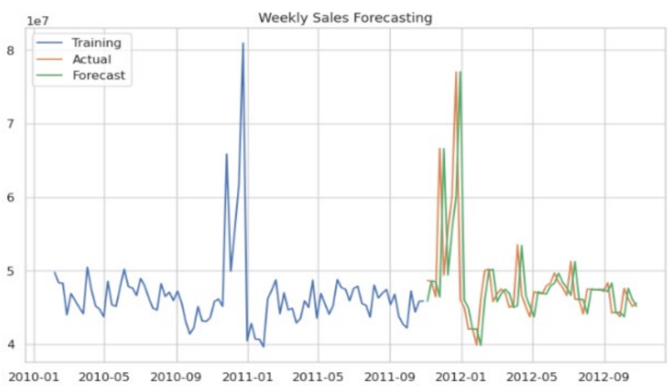

Figure 3. Weekly sales forecast curve (2010-2012).

According to the results presented in Table 1, the Informer model demonstrates significantly superior predictive performance compared to other baseline models in the task of weekly furniture sales forecasting on the Amazon platform. Across the three error-based metrics used for time series forecasting, namely MAE, RMSE and MAPE, the Informer model resulted in lower forecast deviations compared to baseline models TCN, LSTM and GRU. With lower MAE value, the Informer model shows a lower overall prediction deviation and thus overall higher accuracy in long sequence forecasting. The lower RMSE value indicate the Informer model had fewer large errors in the long sequence forecasting, which is critical for sales forecast through seasonality and response to extraneous factors. The superiority of the Informer model is confirmed by the lower MAPE value, i.e. higher accuracy in percentage terms, which is particularly important for e-commerce applications for sales planning and warehouse management. Overall, the result demonstrate that the Informer model could create huge valuation when adapted to the e-commerce sales predications. With more accurate sales forecast, e-commerce sellers and platforms could benefit from having sufficient inventory to capture revenue opportunities while reducing the loss from overstocking and destocking.

Table 1. Forecasting results across models (average of 5 furniture categories).

Model	MAE	RMSE	MAPE (%)
TCN	58.4	73.2	12.6
LSTM	45.7	61.8	9.8
GRU	44.1	60.5	9.5
Informer	**37.3**	**52.7**	**7.6**

6. Conclusions

This study proposes a prediction framework based on the Informer model for addressing the critical task of sales forecasting on e-commerce platforms, with a focus on weekly furniture sales on the Amazon platform. Accurate sales forecasting plays a vital role in optimizing inventory management, enabling dynamic pricing, and supporting personalized marketing strategies. However, the volatility of consumer demand, the influence of macroeconomic indicators and weather conditions, and the interference of promotional events make this task highly challenging. By leveraging sales data of five furniture categories from 2015 to 2020, supplemented with auxiliary features such as price, discount rate, CPI, and weather variables, this study conducted comprehensive data preprocessing, including cleaning, handling missing values, and feature engineering, to construct a high-quality dataset for model training.

Methodologically, this research adopts the Informer model, a Transformer-based architecture that excels in modeling long-sequence time series. Through its encoder–decoder structure, the model effectively captures long-term dependencies while integrating exogenous variables to account for complex nonlinear effects. Experimental comparisons with baseline models including TCN, LSTM, and GRU demonstrate that Informer achieves superior predictive performance. Specifically, the Informer model achieves a Mean Absolute Error (MAE) of 37.3, and a Root Mean Squared Error (RMSE) of 52.7, significantly outperforming alternative approaches. These results suggest that the proposed framework not only accurately reflects long-term sales trends but also captures short-term fluctuations, with forecast curves closely aligned to actual sales dynamics.

The findings of this study provide valuable implications for the operation of e-commerce platforms. In practice, the proposed model can support refined inventory management, helping avoid stockouts caused by underestimation and overstocking caused by overestimation. Moreover, accurate sales forecasting can enhance dynamic pricing mechanisms and improve the targeting of marketing and promotional campaigns, thereby increasing operational efficiency and customer satisfaction.

Despite these promising results, some limitations remain. The current framework relies primarily on historical sales and selected external features, without incorporating consumer behavior data, textual sentiment from reviews, holiday effects, or sudden events. Future research could explore the integration of multimodal data to further enhance predictive accuracy. Additionally, adaptive hyperparameter optimization, ensemble learning, and causal inference methods could be investigated to improve model robustness and interpretability, making the framework more applicable to the highly dynamic and complex environment of e-commerce platforms.

In conclusion, this study, through the application of the Informer model and the integration of multi-source features, demonstrates its superior performance in e-commerce sales forecasting, providing new insights for both algorithmic applications and operational practices in the field.

References

1. Z. B. Yusof, "Analyzing the role of predictive analytics and machine learning techniques in optimizing inventory management and demand forecasting for e-commerce," *International Journal of Applied Machine Learning*, vol. 4, no. 11, pp. 16-31, 2024.
2. P. Ramos, and J. M. Oliveira, "Robust Sales forecasting Using Deep Learning with Static and Dynamic Covariates," *Appl. Syst. Innov. 2023*, vol. 6, p. 85, 2023. doi: 10.20944/preprints202308.0427.v1
3. 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*, vol. 339, no. 1, pp. 297-328, 2024.
4. C. O. Ikeh, "Neural Network Applications in Product Lifecycle Management: Forecasting Feature Adoption and User Retention in Competitive Markets," .
5. Y. Qiao, Z. Zhang, and A. Wilson, "A Lightweight and Interpretable DepGraph-based Pruned Artificial Neural Network for Advertisement Click Prediction," *Advance in Information Technology and Computer Science*, 2024. doi: 10.26855/aitcs.2024.12.006
6. W. Liu, Y. Gu, and Y. Ge, "Multi-factor stock trading strategy based on DQN with multi-BiGRU and multi-head ProbSparse self-attention," *Applied Intelligence*, vol. 54, no. 7, pp. 5417-5440, 2024. doi: 10.1007/s10489-024-05463-5
7. J. Lee, H. J. Kim, and C. Lim, "Long-term forecasting using transformer based on multiple time series," *The Korean Journal of Applied Statistics*, vol. 37, no. 5, pp. 583-598, 2024. doi: 10.5351/kjas.2024.37.5.583
8. X. Li, "Predicting E-Commerce Sales with Three Machine Learning Models," In *2024 2nd International Conference on Management Innovation and Economy Development (MIED 2024)*, October, 2024, pp. 445-454. doi: 10.2991/978-94-6463-542-3_52