

Machine Learning-Based Customer Demand and Online Sales Prediction

Mr. Swapnil Joshi¹

¹ Assistant Professor, Department of Computer Sciences and Applications, Mandsaur University, Mandsaur (M.P.), India

ABSTRACT: The development of Internet technology has changed how customers engage with e-commerce websites, where transactional and customer-related data may offer insightful information on consumer behavior. Companies must handle massive amounts of multivariate data to enhance demand analysis and provide efficient services in response to growing consumer expectations and fierce online competition. The e-commerce sales forecasting problem is addressed in this study, based on actual transactional, temporal, logistical, pricing, payment, and review data from the Olist Brazilian E-Commerce Dataset. The proposed methodology uses data preprocessing techniques such as missing value treatment, date-time transformation, duplicate removal, label encoding, Z-score standardization, and SMOTE-based data balancing. A complex sequential pattern learning model, called Long Short-Term Memory (LSTM) is built to learn complex sequential patterns and predict customer demands and online sales. The experimental results have shown that the proposed LSTM has been able to attain 99.72% accuracy, 99.64% precision, 99.16% recall, and 99.26% F1-score. The results show that the proposed methodology is useful in analyzing e-commerce demand and forecasting e-commerce sales. The model can help businesses manage their inventory, sales, resources, and decision-making processes.

KEYWORDS: Demand Forecasting, Predictive Analytics, Deep Learning, Demand Sensing, Sales Forecasting, Retail Store, E-Commerce.

1. Introduction

The rapid growth of the Internet and digital commerce has transformed the way customers search for products, make purchases, and interact with online businesses. Ecommerce platforms produce enormous quantities of transactional data, such as customer data, product details, order history, pricing information, payment details, reviews, and delivery details [1][2]. This data offer important insights to businesses on how customers are buying and help them anticipate future demand and plan sales accordingly [3]. Estimation of customer demand and online sales is becoming more and more essential since it helps in inventory planning, resource allocation, marketing decisions and effective sales strategy [4][5].

The highly competitive nature of the e-commerce market has increased the importance of understanding customer preferences and purchasing behavior [6]. Past buying experiences, service quality, customer satisfaction, and customer engagement may affect repeat sales and customer relationships [7][8]. Poor inventory management can lead to stockouts,

overstocking, late order delivery, and unhappy customers. Historical data-driven or manual forecasting methods may find it challenging to account for evolving consumer patterns, seasonal trends, price influence, and intricate market dynamics [9][10].

However, the ease of access to vast historical e-commerce data opens the door to the creation of predictive systems that can uncover significant relationships between customer behavior, product offerings, transactions, and sales results [11]. Businesses can leverage historical buying trends to predict future demand, manage inventory better, make smarter promotional offers and make better operational decisions [12][13][14]. Based on the above, transactional, temporal, price, review, and customer-related attributes can be used to provide complementary information to understand online sales behavior [15][16]. Accordingly, customer demand and forecasting have been an active research field to support of intelligent and scalable decision-making in the modern e-commerce environment [17].

To address these challenges, Machine Learning (ML)

and Deep Learning (DL) have emerged as more effective techniques to analyze complex, high-dimensional e-commerce data. The ML approaches can find nonlinear relations and concealed patterns, and the DL architectures can learn complex temporal and sequential dependencies from massive amounts of data [18][19]. Accordingly, this research suggests a framework of Customer Demand and Online Sales Prediction based on Machine Learning and Deep Learning techniques to analyze historical e-commerce data and make accurate predictions for effective sales planning and smart business decisions.

A. Motivation and Contribution

The motivation for this study is the fact that as a result of the growing amount of information about the transactions of e-commerce, data accuracy in this area is crucial if the need to accurately predict customer demand and online sales are considered. The volatility of buying trends, seasonal trends, price changes, and complex customer dynamics are not reflected by many traditional forecasting methods. Therefore, a smart predictive model must be created and examine historical e-commerce data to identify meaningful demand patterns. The proposed solution will facilitate sales planning, inventory management, resource allocation and data-driven business decisions. This research offers several key contributions as listed below:

- Using the Olist Brazilian E-Commerce Dataset, created an LSTM model to analyze consumer demand and forecast online sales.
- Used SMOTE to address class imbalance and improve the representation of minority-class samples.
- Demonstrated that the proposed LSTM can support sales forecasting, demand planning, inventory management, and data-driven e-commerce decision-making.
- Used F1-score, recall, accuracy, and precision to assess the suggested model.
- Analyzed customer, product, payment, review, seller, and delivery information to identify important patterns and relationships in e-commerce transactions.
- Provided a reliable framework for improving customer demand estimation and online sales prediction in modern e-commerce environments.

The study is justified by the need for accurate customer demand and online sales prediction to support effective inventory management and business planning. The proposed LSTM model is novel in its ability to capture sequential patterns from integrated

transactional, temporal, pricing, payment, review, and delivery data. The methodology includes data preprocessing, feature transformation, standardization, and class balancing, which enhance the data quality and predictability of the classification results. The proposed method offers an effective solution for the analysis of demands and data-driven decision-making in e-commerce.

B. Paper Organization

The remaining sections of the document are structured as follows: The literature review is included in Section II. The research methodology is covered in Section III, the case study and discussion of results are in Section IV, and the study is concluded in Section V.

2. Literature Review

A systematic review of research on Customer Demand Analysis and Online Sales Forecasting was conducted to gain insights into current methods, data and predictive approaches in this area. The outcome of these studies is summarized concisely in Table I, which identifies the models used and the datasets available, the main outcome, and the limitations.

Singh (2026) predicts future sales for the Brazilian online business Olist based on transactional, logistical, price, and review data. Soon after the timestamps were normalized, records lacking delivery fields were removed, and data was aggregated daily. Then, features based on time, delayed time, rolling data, and trends were created. Prior to their combination through weighted ensemble fusion, XGBoost, LightGBM, and Extra Trees were each improved using RandomizedSearchCV. While individual models and classical baselines performed poorly, the ensemble outperformed them with an MAE of 7.1086, RMSE of 9.7559, and R^3 of 0.9820. When it comes to nonlinear e-commerce demand forecasting, lag and rolling features are crucial, and the ablation findings demonstrated that full-time features performed the best (MAE 6.41, RMSE 8.23, R^2 0.9872) [20].

P.Vanitha et al. (2026) proposed framework integrates feature normalization, latent customer behavior encoding, campaign influence modeling, and temporal sales prediction into a unified optimization framework. Promotion-aware inputs support predictive learning, while marketing decisions use constrained revenue maximization. The model achieved a revenue forecasting RMSE of 0.071, mean absolute sales error of 0.052, and variance explanation ratio of 0.942. It also increased campaign response probability by 28.4%, achieved a revenue/marketing cost ratio of 3.62, and 98.1% budget utilization efficiency [21].

Shobha et al. (2025) offer issues including low

profitability, low average orders, high operational expenses, and last-mile delivery inefficiencies. Proposing a new operating model that uses machine learning to decrease operational and last-mile delivery expenses is the plan. they suggest a DDFD, or Dynamic Demand Forecast Dashboard, that can predict purchases for certain categories at specific times and places. Authoritative integrated moving average (ARIMA) models and the Random Forest classifier do this with a 70% success rate. In addition to maximizing income in real time, the dynamic pricing strategy may adjust to changes in the market accordingly [22].

Shrivastava (2025) presents a comprehensive analysis of AI and machine learning methods for predicting return rates in online retail using behavioral and transactional data. Literature review and empirical testing on 6,600 customer transactions identified key predictive features and compared machine learning methods. Ensemble methods, particularly XGBoost and CatBoost, achieved the best results with F1-scores above 0.92, while session time, previous return history, and transaction patterns were important predictors [23].

Iqbal et al. (2024) the goal of product rating system is to convince buyers to buy any goods. Success in the marketplace is another indicator of a product's quality. This study employed a machine learning model to predict a product's rating according to its category, stated price, discounted price, and quantity of reviews. For this investigation, they used the Amazon dataset that is available on Kaggle. In order to forecast product

ratings depending on category, stated price, discounted price, and number of reviews, this research uses an Amazon dataset with 61,950 electrical appliance items to assess several machine learning models. With a recall of 0.6 and an accuracy of 0.65 for 1-star ratings, the Random Forest model shows greater performance compared to others. It also outperforms for 2-star, 4-star, and 5-star forecasts [24].

Almadhor et al. (2023) proposes an ensemble voting method based on Term Frequency-Inverse Document Frequency (TF-IDF) features for classifying mobile phone ratings using over 13,000 Flipkart smartphone reviews. Data cleansing, balancing, feature extraction using TF-IDF, and prediction based on voting are all part of the technique. Experimental results showed improved accuracy and efficiency over conventional machine learning methods, achieving a maximum accuracy of 98.0%. The method can be extended to other e-commerce platforms with large online review datasets to support informed purchase decisions [25].

K et al. (2023) a sentiment analysis approach using the "Amazon and Yelp Reviews" dataset through daily data collection, comment preprocessing, and a BiLSTM model. Quantitative measures of customer satisfaction include F1-score, recall, accuracy, and precision. Similar to other deep learning approaches like deep transfer learning and character-level CNN-LSTM, the suggested method obtains an average accuracy of 87.3%. The approach can support customer feedback analysis, product and service development, and online reputation management [26].

TABLE I. Recent Studies on Customer Demand Analysis and Online Sales Forecasting

Author	Proposed Work	Dataset	Key Findings	Limitations & Future Work
Singh (2026)	Ensemble-based e-commerce sales forecasting using XGBoost, LightGBM, and Extra Trees with temporal, lag, rolling, and trend features.	Olist Brazilian E-Commerce Dataset	Weighted ensemble achieved MAE 7.1086, RMSE 9.7559, R ² 0.9820; time-based features improved performance.	Focuses mainly on sales forecasting; future work can explore deep learning and broader customer-demand behavior.
P. Vanitha et al. (2026)	Integrated framework for customer behavior encoding, campaign influence modeling, and temporal sales prediction.	E-commerce sales/marketing dataset	Achieved RMSE 0.071, MAE 0.052, R ² 0.942, with improved campaign response and marketing efficiency.	Greater adaptability across different e-commerce platforms and real-time deployment can be investigated.
Shobha et al. (2025)	Dynamic Demand Forecast Dashboard using ARIMA, Random Forest, and dynamic pricing.	E-commerce sales data	The proposed system achieved 70% accuracy and supported location-, category-, and time-based demand forecasting.	Relatively limited predictive accuracy; future work can employ advanced deep learning and larger datasets.
Shrivastava (2025)	AI/ML-based prediction of online retail product returns	6,600 customer transactions	XGBoost and CatBoost achieved F1-scores above 0.92; behavioral	Concentrates on return prediction; future research can address integrated sales and demand forecasting.

	using behavioral and transactional features.		and transaction features were important.	
Iqbal et al. (2024)	Machine learning-based product-rating prediction using product and pricing attributes.	61,950 Amazon electronics records	Random Forest showed superior performance for several product-rating categories.	Product-rating prediction is different from demand and sales forecasting; additional temporal and customer features are needed.
Almadhor et al. (2023)	TF-IDF-based ensemble voting model for classifying smartphone reviews.	13,000+ Flipkart reviews	Ensemble voting achieved 98.0% accuracy for review classification.	Primarily sentiment/review classification; future work can connect customer sentiment with demand and sales forecasting.
K et al. (2023)	BiLSTM-based sentiment analysis of customer reviews.	Amazon and Yelp Reviews Dataset	BiLSTM achieved an average 87.3% accuracy for customer sentiment analysis.	Focuses on sentiment rather than demand forecasting; additional sales and temporal variables could improve business prediction.

Research gaps: Existing studies mainly focus on sales forecasting, marketing performance, product ratings, reviews, or return prediction using individual machine learning and ensemble methods. Limited research jointly analyzes customer demand and online sales using integrated transactional, temporal, pricing,

payment, review, and delivery information. Conventional approaches may also struggle to capture complex sequential purchasing patterns and changing customer behaviour. Therefore, an LSTM-based framework is required to effectively learn temporal dependencies and improve customer demand and online sales prediction.

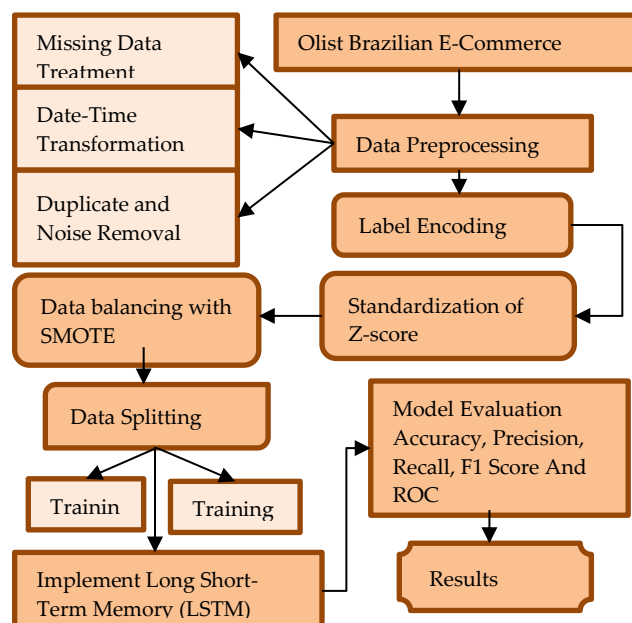


Fig. 1. Proposed Framework for Customer Demand Analysis and Online Sales Forecasting

3. Methodology

The methodology includes data acquisition from the Olist Brazilian E-Commerce Dataset, followed by data cleaning, missing-value treatment, date-time transformation, duplicate removal, label encoding, Z-score standardization, and SMOTE-based class balancing. The processed data was split 75:25 across the training, validation, and testing sets. An LSTM model was then trained to forecast online sales and customer demand. The accuracy, precision, recall, F1-score,

confusion matrix, and ROC-AUC were used to assess the model's performance. Fig. 1 shows the proposed flowchart of Customer Demand Analysis and Online Sales Forecasting.

A. Data Acquisition and Analysis

The Olist Brazilian E-Commerce Dataset consists of 9 interconnected datasets, with about 111,468 orders ranging from 2016 to 2018, including information on customers, products, sellers, payments, reviews, and deliveries. The sales distribution, customer demand,

purchasing patterns, and correlations between features were studied using data visualizations like bar plots and heatmaps, as shown below:

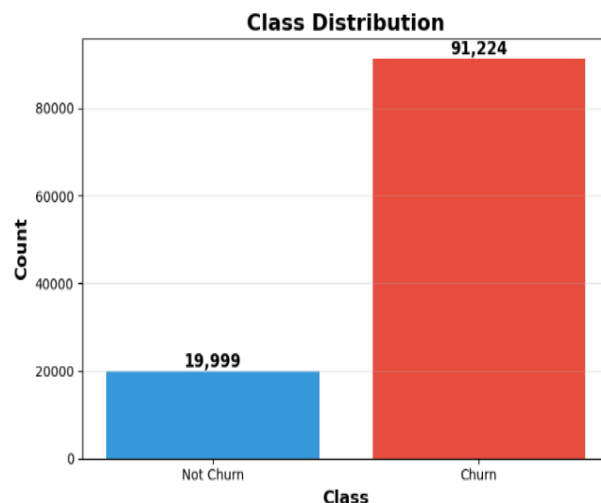


Fig. 2. Bar graph of class distribution

The initial class distribution of the dataset is shown in Fig. 2. There are 19,999 instances in the Not Churn class and 91,224 instances in the Churn class. This

considerable difference indicates a strong class imbalance. This imbalance may lead to the majority class being favored in training of the predictive model.

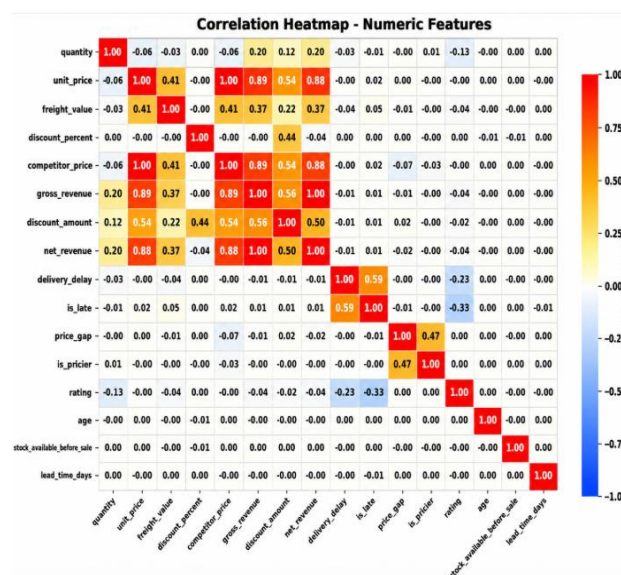


Fig. 3. Correlation Matrix Heatmap

The heatmap in Fig. 3 indicates strong positive correlations among numerical features, such as unit price, competitor price, gross revenue and net revenue [27]. The majority of other features are poorly correlated, and delivery delay and rating are moderately negatively correlated with the variables related to lateness.

B. Data Pre-Processing

The Olist Brazilian E-Commerce Dataset underwent data integration, cleansing, and feature engineering to ensure its readiness for model development. The preprocessing involved Missing Data Treatment, Date-Time Transformation, Duplicate and Noise Removal, Label Encoding, and Data Normalization. The most important preprocessing steps are summarized as

follows:

- **Missing Data Treatment:** Missing or incomplete data were identified and treated appropriately considering the nature of the variables. This facilitated the consistency of the data and reduced the effect of missing information on models' performance.
- **Date-Time Transformation:** Date related attributes that were relevant were transformed to a standardized format of datetime. The Olist Online Dataset was queried to isolate the important temporal patterns of day, month, hour, minute and second.
- **Duplicate and Noise Removal:** Duplicate records, inconsistent values and irrelevant observations were

identified and removed. This minimized data redundancy and enhanced the confidence of the data for subsequent analysis.

- **Label Encoding:** Categorical variables were encoded into numerical labels for use in machine learning models. The categorical information of the original dataset was retained by assigning each category a unique integer value.

C. Standardization of Z-score

The data was standardized using Z-scores to keep the values equal among the features and minimize the dominance of any feature. Every feature was normalized by subtracting its mean value and dividing by its Standard Deviation. This process normalized all the numerical attributes, which helped the model to converge and be trained in a more stable way. The Equation (1), which is used to standardize values to Z-scores:

$$Z = \frac{X - \mu}{\sigma} \quad (1)$$

- Z = Standardized Value
- X = Original Value
- μ = mean of the feature
- σ = Standard Deviation of the feature

This transformation ensures that the mean of the data is zero and the Standard Deviation is one.

D. Data balancing using Synthetic Minority Over-sampling Technique (SMOTE)

Data balancing is a preprocessing method that evens out sample distributions across classes, which helps with class imbalance. This research used SMOTE to create artificial samples from the underrepresented group. SMOTE identifies the nearest neighbors of each minority-class sample using Euclidean distance and generates new samples through interpolation between the selected sample and its neighbors. The newly generated instances are added to the training dataset, increasing minority-class representation and producing a more balanced class distribution for model development.

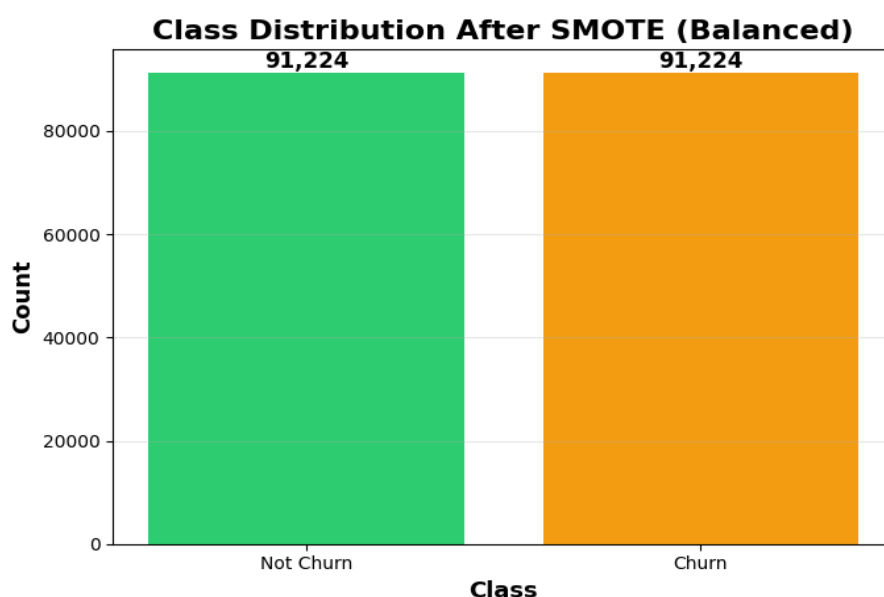


Fig. 4. Bar graph of class distribution after SMOTE

The class distribution following SMOTE application is shown in Fig. 4. The dataset is evenly distributed between the Not Churn and Churn classes, with 91,224 occurrences in each. To help reduce class disparities, SMOTE creates synthetic samples for underrepresented groups. The balanced data can enhance the learning and classification capability of the model for both classes.

E. Data Splitting

The dataset Olist Online was split into a training set and a testing set with the proportions of 75:25. The training set was divided into a basic training set and a validation set after standardization. This enables successful model training, hyperparameter tuning, and

impartial assessment.

F. Proposed Long Short-Term Memory (LSTM) Model

This study utilizes a Long Short-Term Memory (LSTM) model that is based on DL to forecast online sales and consumer demand. An LSTM is a kind of RNN that takes outputs from earlier layers and learns their dependencies over time. It can process sequential data thanks to its feedback links. The memory cell in an LSTM node maintains values throughout time while the gates control the flow of information. The node also has an input gate, an output gate, and a forget gate. Recurrently linked memory blocks with three multiplicative gates control information for a predetermined duration in the

LSTM layers through continuous write, read, and reset operations. It is as follows in Equations (2) to (7) that the unit's input and output are updated:

Gates:

$$i_t = \sigma(W_i x_t + U_i h_{t-1} + b_i) \quad (2)$$

$$f_t = \sigma(W_f x_t + U_f h_{t-1} + b_f) \quad (3)$$

$$o_t = \sigma(W_o x_t + U_o h_{t-1} + b_o) \quad (4)$$

Input transform:

$$g_t = \tanh(W_g x_t + U_g h_{t-1} + b_g) \quad (5)$$

State update:

$$c_t = f_t \odot c_{t-1} + i_t \odot g_t \quad (6)$$

$$h_t = o_t \odot \tanh(c_t) \quad (7)$$

The logistic sigmoid function and element-wise multiplication are represented by σ and $\odot$, respectively, in the equations given above. With each time step t , the LSTM unit has the following components: an input gate i_t , a forget gate f_t , an output gate o_t , a hidden unit h_t , and a memory cell c_t . The learnt parameters are W and U , while the extra bias is denoted by b . The input gate seems to regulate the amount of data written to each unit, the forget gate the amount of data wiped from the memory cell, and the output gate the amount of data shown from the internal memory. The suggested LSTM model was set up with the following parameters: 20 epochs, 32 batches, 0.001 learning rate, and Adam optimizer. For efficient learning, the model made use of 128 LSTM units, a dropout rate of 0.2, and ReLU activation. A sigmoid output layer was employed for classification with a binary cross-entropy loss function.

G. Evaluation Metrics

The suggested model was evaluated using a number of industry-standard measures. The True Negatives (TN), False Negatives (FN), True Positives (TP), and False Positives (FP) were calculated by representing the true and false predictions for each class in a confusion

matrix. In order to conduct a thorough evaluation of the model, the following metrics were used: accuracy, precision, recall, and F1-score, as shown in Equations (8) to (11):

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (8)$$

$$Precision = \frac{TP}{TP+FP} \quad (9)$$

$$Recall = \frac{TP}{TP+FN} \quad (10)$$

$$F1 - score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \quad (11)$$

The accuracy is the percentage of the data set that is correctly predicted. The precision is the ratio of true positives to the total number of positive predictions. Recall is the percentage of true positive responses that the model recognizes. F1-score is the harmonic average of both precision and recall, thus showing a balanced classification measure. The ROC Curve shows the tradeoff between the TPR and FPR for different thresholds of classification, with TPR being equivalent to recall and FPR being equivalent to 1-specificity.

4. Results and Discussion

This section outlines the experimental setup and the performance of the proposed model in customer demand and online sales prediction, highlighting its ability to evaluate and process efficiently. The tests were executed on a state-of-the-art desktop PC that ran Windows 11 Pro and included a 3.0 GHz Intel Core i9-13900K CPU, 64 GB of DDR5 RAM, and a 24 GB VRAM NVIDIA RTX 4090 GPU. Pandas, NumPy, Matplotlib, Seaborn, and Scikit-learn were among the Python libraries used for the implementation and analysis. With a 99.72% accuracy rate, 99.64% precision rate, 99.16% recall rate, and 99.26% F1-score, the suggested Long Short-Term Memory (LSTM) model performed well in classification, as shown in Table II. These results demonstrate that the proposed deep learning-based LSTM model provides highly effective and reliable performance for customer demand and online sales prediction.

TABLE II. Classification Performance of the Proposed Predictive Analysis Model for Customer Demand and Online Sales

Matrix	Proposed Long Short-Term Memory (LSTM) Model
Accuracy	99.72
Precision	99.64
Recall	99.16
F1-score	99.26

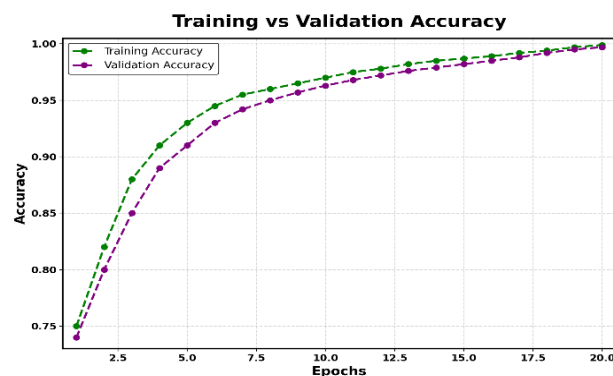


Fig. 5. LSTM Training and Validation Accuracy

The accuracy of the proposed LSTM model during training and validation is shown in Fig. 5, spanning 20 epochs. As the number of epochs rises, both curves demonstrate a constant improvement in accuracy. The

training accuracy reaches nearly 100%, while the validation accuracy also approaches 99.7%. The similarity level of the two curves illustrates good generalization and stable learning, with few signs of overfitting.



Fig. 6. LSTM Training vs Validation Loss

The training and validation loss over the 20 epochs are presented in Fig. 6. Both of the losses are monotonically decreasing and LSTM model is converging, showing good learning. The training and

validation curves remain close to each other, suggesting good generalization with minimal overfitting. The resulting losses are small at the final epochs, indicating stable and accurate model training.

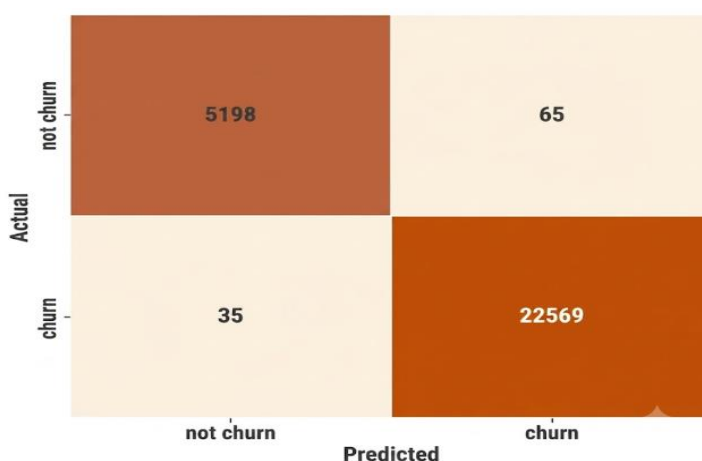


Fig. 7. Confusion Matrix for LSTM Model

The classification accuracy of the customer churn across 27,867 instances is shown in Fig. 7, where 22,569 churn cases were successfully predicted and 5,198 cases of non-churn were successfully predicted. The model

made 65 false positive and 35 false negatives, showing the high accuracy and reliability of the classification. Overall results reveal that the model performs well in terms of correct classification of churned and non-churned customers as well as minimizing misclassification.

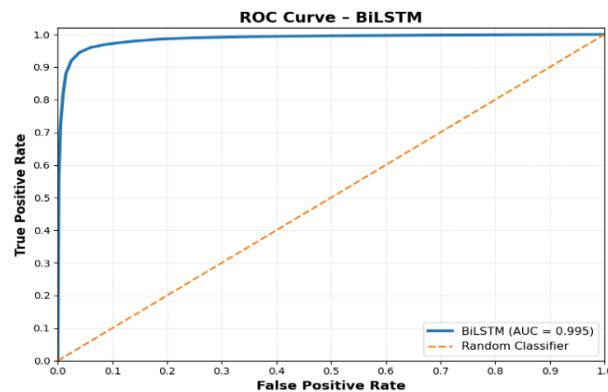


Fig. 7. ROC Curve of the LSTM Model

The classification accuracy of the proposed LSTM model for churn vs non-churn customers is presented in Fig. 8. The model achieves an excellent AUC of 0.995 (99.5%), indicating a very high ability to correctly discriminate between the two classes. The ROC curve remains close to the upper-left corner, demonstrating a high true-positive rate with a very low false-positive rate.

A. Comparative Analysis

A comparison was made between the suggested

Long Short-Term Memory (LSTM) model and XGBoost, Decision Tree (DT), Support Vector Machine (SVM), and Convolutional Neural Network (CNN) models to assess their performance. The results are shown in Table III. With a maximum accuracy of 99.72%, precision of 99.64%, and recall of 99.16%, the suggested LSTM proved to be the most predictive. Additionally, it attained an F1-score of 99.26%, suggesting a robust equilibrium between recall and accuracy. In terms of online sales and consumer demand prediction, the findings show that the suggested LSTM model is superior to the comparative deep learning and machine learning methods.

TABLE III. Comparative Evaluation of Machine Learning and Deep Learning Approaches for Customer Demand and Online Sales Prediction

Model	Accuracy	Precision	Recall	F1-score
XGBoost[28]	81	79	83	81
DT[29]	84	71	73	84
SVM[30]	87	98.7	21.5	93
CNN[28]	97.35	97.42	93.62	97.51
Proposed LSTM	99.72	99.64	99.16	99.26

The proposed Long Short-Term Memory (LSTM) model is more accurate in forecasting customer demand and online sales, reaching an accuracy of 99.72%. It has the capability of learning long-term temporal dependencies, which means that it can be used to model the complicated purchasing and sales patterns accurately. The model also provides high precision, recall, and F1-score, ensuring reliable and balanced predictions. Thus, the proposed LSTM has good potential for accurate and efficient online sales and customer demand forecasting.

5. Conclusion and Future Study

The rise of artificial intelligence, machine learning, and deep learning opens a tremendous potential for customer demand analysis and online sales prediction in e-commerce. In this study, the Olist Brazilian E-Commerce dataset was analyzed and a Long Short-

Term Memory (LSTM) model was proposed to determine customer demand patterns and predict online sales. The proposed LSTM model has demonstrated an excellent accuracy rate of 99.72%, which is higher than the other models such as XGBoost (81%), Decision Tree (84%), SVM (87%), and CNN (97.35%). The data shows that LSTM is capable of learning sequential relationships and complex purchasing patterns in e-commerce data. The suggested solution can enable demand forecasting, sales plans, inventory management and resourcing. The results overall validate that LSTM is a very effective and reliable method for customer demand analysis and online sales forecasting, which can facilitate intelligent and data-driven decision-making in modern e-commerce environments.

Future work can overcome the shortcomings of processing the single e-commerce dataset and historical transaction data by using larger real-time datasets and

including more customer and market data and advanced hybrid deep learning models. Other studies can look into real-time forecasting and deployment to help with dynamic inventory and sales.

REFERENCES

1. Z. H. Kilimci et al., "An Improved Demand Forecasting Model Using Deep Learning Approach and Proposed Decision Integration Strategy for Supply Chain," *Complexity*, vol. 2019, no. 1, Jan. 2019, doi: 10.1155/2019/9067367.
2. S. Sreevathsa, "Bridging Technical Telemetry and Business Decision-Making through Data Product Management," *Univers. Libr. Innov. Res. Stud.*, vol. 3, no. 3, pp. 32–36, Aug. 2026, doi: 10.70315/uloap.ulirs.2026.0303006.
3. R. Palwe, "The Role of User Experience Design in Advancing Digital Sustainability," *Int. J. Sci. Res.*, vol. 14, no. 10, pp. 1168–1176, 2025.
4. Y. Ahmadov and P. Helo, "Deep learning-based approach for forecasting intermittent online sales," *Discov. Artif. Intell.*, vol. 3, no. 1, p. 45, Dec. 2023, doi: 10.1007/s44163-023-00085-1.
5. R. S. Ahmed, M. Hasnain, M. H. Mahmood, and M. A. Mehmood, "Comparison of Deep Learning Algorithms for Retail Sales Forecasting," *ICCK Trans. Intell. Syst.*, vol. 1, no. 3, pp. 112–126, Oct. 2024, doi: 10.62762/TIS.2024.300700.
6. R. Acharya, D. P. Sahu, J. Singh, R. Tomar, A. P. Singh, and Neetu, "An Effective Machine Learning Model for Accurate ROI Prediction and Optimization in Marketing Strategies," in *2025 International Conference on Technology Enabled Economic Changes (InTech)*, IEEE, Feb. 2025, pp. 520–526. doi: 10.1109/InTech64186.2025.11198202.
7. K. Alice, S. H. ul H. Andrabi, and S. Jha, "Sales Forecasting Based on Ensemble Learning," Aug. 2023. doi: 10.36227/techrxiv.24049452.
8. S. Bauskar, "Predictive Analytics for Sales Forecasting in Enterprise Resource Planning (ERP) Systems Using Machine Learning Technique," *Int. Res. J. Mod. Eng. Technol. Sci.*, vol. 04, no. 06, pp. 2582–5208, 2022, doi: 10.56726/IRJMETS26271.
9. S. S. Yadav, "Sales Forecasting for Small Medium Enterprises Using Machine Learning," *National College of Ireland*, 2023.
10. B. Biswas, M. K. Sanyal, and T. Mukherjee, "AI-Based Sales Forecasting Model for Digital Marketing," *Int. J. E-bus. Res.*, vol. 19, no. 1, pp. 1–14, Feb. 2023, doi: 10.4018/IJEER.317888.
11. V. Chaturvedi, R. Narra, and S. K. Chintagunta, *Applied AI Engineering for Developers: Building Intelligent Applications at Scale*. Wissira Research Lab, 2026. doi: 10.63345/WP-978-93-7559-963-0.
12. T. de C. Moraes, X. Yuan, and E. P. Chew, "Hybrid convolutional long short-term memory models for sales forecasting in retail," *J. Forecast.*, vol. 43, no. 5, pp. 1278–1293, Aug. 2024, doi: 10.1002/for.3073.
13. V. Zalizniuk and O. Nozhovnik, "Express Method for Calculating Gross Margin in E-Commerce: A Practical Approach," *Balt. J. Econ. Stud.*, vol. 10, no. 4, pp. 201–210, Dec. 2024, doi: 10.30525/2256-0742/2024-10-4-201-210.
14. Ö. Büyükakın and Ö. Zeybek, "Share of Local Assortments in Revenue for a Large Scale National Food Retailer : Assessment of Long and Short Time Series Patterns," 2024, pp. 57–63. doi: 10.1007/978-3-031-67195-1_8.
15. S. Singh, S. A. Pahune, P. Chatterjee, and R. Sura, "Advanced Machine Learning Techniques for Prediction of Customer Churn in Telecommunication Sector," in *2025 IEEE 6th India Council International Subsections Conference (INDISCON)*, Rourkela, India: IEEE, 2025, pp. 1–6, August. doi: 10.1109/INDISCON66021.2025.11252233.
16. S. Kukkar, "Leveraging Price-Quote Data with AI for Smarter Configure-Price-Quote Systems," in *2026 2nd International Conference on Artificial Intelligence Systems (AIS)*, San Antonio, TX, USA: IEEE, 2026, pp. 329–335, May. doi: 10.1109/AIS69919.2026.11621526.
17. E. Aldhahri, "The Use of Recurrent Nets for the Prediction of e-Commerce Sales," *Eng. Technol. Appl. Sci. Res.*, vol. 13, no. 3, pp. 10931–10935, Jun. 2023, doi: 10.48084/etasr.5964.
18. C.-J. Liu, T.-S. Huang, P.-T. Ho, J.-C. Huang, and C.-T. Hsieh, "Machine learning-based e-commerce platform repurchase customer prediction model," *PLoS One*, vol. 15, no. 12, p. e0243105, Dec. 2020, doi: 10.1371/journal.pone.0243105.
19. H. Huang et al., "A Cross-Platform Consumer Behavior Analysis of Large-Scale Mobile Shopping Data," in *Proceedings of the 2018 World Wide Web Conference on World Wide Web - WWW '18*, New York, New York, USA: ACM Press, 2018, pp. 1785–1794. doi: 10.1145/3178876.3186169.
20. N. Singh, "Advanced Machine Learning Techniques for Enhancing Accuracy in E-Commerce Sales Forecasting," in *2026 International Conference on Modern Computing, Networking and Applications*

- (MCNA), IEEE, May 2026, pp. 115–122. doi: 10.1109/MCNA69896.2026.11574094.
21. P. Vanitha, S. S. N. S, R. T, and P. V, "Data-Driven Marketing and Sales Prediction for Small and Medium E-Commerce Companies," in 2026 IEEE Nexo-Tech - International Conference on Advanced Technologies and Innovations (ICATI), IEEE, May 2026, pp. 1–6. doi: 10.1109/ICATI69399.2026.11604384.
22. R. Shobha, S. S. Selvi, M. S. Chowdary, P. N, V. N, and H. K. Omkumar, "Dynamic Demand Forecast Dashboard – A Machine Learning based Quick Commerce Model," in 2025 International Conference on Intelligent and Innovative Technologies in Computing, Electrical and Electronics (IITCEE), IEEE, Jan. 2025, pp. 1–6. doi: 10.1109/IITCEE64140.2025.10915287.
23. S. Shrivastava, "AI-Driven Return Rate Prediction in Online Retail using Behavioral and Transactional Data," in 2025 3rd International Conference on Sustainable Computing and Data Communication Systems (ICSCDS), IEEE, Aug. 2025, pp. 932–937. doi: 10.1109/ICSCDS65426.2025.11166988.
24. M. Iqbal, F. Hamayat, R. F. Ahmad, and S. K. Husnain, "A Machine Learning-Based Framework for Predicting E-Commerce Product Ratings," in Proceedings of 2024 21st International Bhurban Conference on Applied Sciences and Technology, IBCAST 2024, 2024. doi: 10.1109/IBCAST61650.2024.10876934.
25. A. Almadhor, S. Abbas, G. A. Sampedro, M. Abisado, and T. R. Gadekallu, "A Novel Smartphone Recommendation System Using Ensemble Machine Learning," in 2023 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), IEEE, Dec. 2023, pp. 1–7. doi: 10.1109/CSDE59766.2023.10487679.
26. G. K, S. K., A. B. J, P. V. L. R. N. S. Sudha, and R. A. C, "Deep Learning and Sentiment Analysis Improve E-commerce Sales Prediction," in 2023 International Conference on Data Science and Network Security (ICDSNS), IEEE, Jul. 2023, pp. 1–7. doi: 10.1109/ICDSNS58469.2023.10245916.
27. D. Patel, "Integrating Price Elasticity and Reinforcement Learning: A Data-Driven Framework for Strategic E-commerce Pricing," in 2026 IEEE 5th International Conference on AI in Cybersecurity (ICAIC), IEEE, Feb. 2026, pp. 1–6. doi: 10.1109/ICAIC67076.2026.11395747.
28. M. Zeinali, L. Ramezani Asli, and M. A. Khalili, "Integrating Business Intelligence and CRM Systems With a Machine Learning Approach for Predictive Customer Retention in E-Commerce," *Sci. World J.*, vol. 2026, no. 1, 2026, doi: 10.1155/tswj/1946904.
29. M. M. Bachir, F. Z. Laallam, M. Mezati, A. Rayane, and D. Ahlam, "Enhancing Customer Satisfaction," *Brazilian J. Technol.*, vol. 7, no. 4, p. e74444, Nov. 2024, doi: 10.38152/bjtv7n4-011.
30. A. N. Wong and B. P. Marikannan, "Optimising e-commerce customer satisfaction with machine learning," *J. Phys. Conf. Ser.*, vol. 1712, no. 1, 2020, doi: 10.1088/1742-6596/1712/1/012044.