International Journal of Modern Computer Science and IT Innovations (IJMCSIT)
eISSN: 3087-4289
Volume03 Issue07 | Page No: - 74-82

Received: 30 May 2026, Revised: 16 June 2026, Accepted: 28 June 2026, Published: 24 July 2026

DOI: https://doi.org/10.55640/ijmcsit-v03i07-07

Deep Learning-Based Customer Segmentation for Targeted Marketing in E-Commerce Platforms

Mr. Swapnil Joshi [1]

[1] Assistant Professor, Department of Computer Sciences and Applications, Mandsaur University, Mandsaur, India

ABSTRACT: The rapid growth of e-commerce platforms has generated massive volumes of customer transaction data, making intelligent customer segmentation essential for personalized marketing and customer relationship management. Traditional segmentation techniques often fail to capture complex and sequential purchasing behaviours, resulting in less effective marketing decisions. This paper presents a deep learning framework for consumer segmentation using the Online Retail dataset from the UCI Machine Learning Repository. The approach is built on Long Short-Term Memory (LSTM). The proposed methodology includes comprehensive data preprocessing, label encoding, RFM (Recency, Frequency, Monetary) feature engineering, Min-Max feature scaling, and supervised LSTM model training. The framework classifies customers into meaningful behavioural segments, including High-Value, Loyal, Potential, Occasional, and At-Risk customers, enabling businesses to implement targeted marketing strategies. The proposed model demonstrates excellent predictive performance, achieving 99.77% accuracy (ACC), 99.39% precision (PRE), 99.93% recall (REC), 99.66% F1-score (F1), and 99.94% ROC-AUC. Comparative analysis with existing customer segmentation approaches confirms the superiority of the proposed LSTM framework in learning customer purchasing patterns. The findings indicate that the proposed approach provides an effective and reliable solution for intelligent customer segmentation, personalized recommendations, and enhanced marketing decision-making in modern e-commerce platforms.

KEYWORDS: Customer Segmentation, Deep Learning, Long Short-Term Memory (LSTM), E-Commerce, Targeted Marketing.

1. Introduction

Market competition has gotten tougher as e-commerce platforms have grown quickly [1]. Businesses are always trying to get and keep customers. The rise of mobile applications, individualised shopping experiences, and ubiquitous internet connections have all contributed to online shopping's status as a cornerstone of contemporary business. Online markets come in many shapes and sizes, from generalised services like Amazon and eBay to niche sites catering to specific niches like food, clothing, and technology [2][3]. As the volume of customer transaction data continues to grow, customer segmentation has become a critical strategy for understanding purchasing behaviour, identifying valuable customer groups, and delivering personalized marketing campaigns [4]. It is crucial to understand client purchase habits through efficient segmentation to improve customer retention, customer happiness, and business sustainability in the face of evolving consumer preferences and behaviours [5][6].

Effective marketing relies on customer segmentation. It aids companies in focusing on certain subsets of consumers that share characteristics [7]. Segmentation used to be based on demographic variables like income, geography, or age. Customers' buying habits in contemporary e-commerce settings are too fluid and ever-changing for them to accurately record [8]. New forms of advertising have emerged. These days, it employs sophisticated analytics and AI. The use of these instruments transformed the field of PRE marketing and consumer segmentation. Segmenting customers helps modern marketers learn more about different demographics. It aims for characteristics motivated by actions rather than generalisations [9]. Marketers may enhance their consumer profiles through data-driven segmentation, allowing for highly personalised marketing efforts.

Ineffective marketing tactics and subpar client involvement are common outcomes of using traditional rule-based segmentation approaches and simple statistical models, which fail to capture the complexities of consumer behaviours [10]. ML and DL techniques have become powerful data-driven ways to divide customers into groups. They do this by looking at a lot of transactional data and finding hidden patterns in how people behave [11][12]. ML algorithms improve segmentation by learning complex relationships among customer attributes, while DL models further enhance performance through automatic feature extraction and the ability to capture nonlinear and sequential dependencies in customer interactions [13][14]. These techniques enable businesses to generate more accurate customer profiles, predict purchasing behaviour, and deliver personalized marketing strategies. Consequently, the adoption of ML and DL has become increasingly important for intelligent customer analytics and data-driven decision-making in modern e-commerce platforms.

A. Motivation and Contributions of the Study

The rising need for smart consumer segmentation methods that permit targeted and personalised marketing initiatives, spurred by the meteoric expansion of e-commerce platforms, is what ultimately led to the decision to conduct this research. Because consumer buying behaviour is both sequential and complicated, traditional techniques of client segmentation, including demographic analysis, rule-based approaches, and traditional ML algorithms frequently fall short. Online marketplaces produce vast amounts of transaction data, which is driving demand for sophisticated DL models that can understand customers' buying habits over time and improve the ACC of consumer classification. Consequently, in order to bolster data-driven targeted marketing decisions and enhance customer segmentation effectiveness, this study suggests an LSTM-based customer segmentation. Below is a summary of the main points made by this study:

- Utilized the Online Retail dataset containing 541,909 real-world e-commerce transaction records.
- Performed data preprocessing, label encoding, RFM feature engineering, and Min–Max scaling to generate high-quality behavioural features.
- Developed an LSTM-based DL framework to learn sequential purchasing patterns for customer segmentation.
- Classified customers into High-Value, Loyal, Potential, Occasional, and At-Risk segments to support targeted marketing.
- Confusion Matrix, ROC-AUC, ACC, PRE, REC, and F1 were used to validate the suggested framework, which outperformed the conventional approaches of consumer segmentation.

B. Justification and Novelty of the Paper

This study justifies the application of DL for customer segmentation. Existing customer segmentation techniques primarily rely on clustering algorithms, conventional ML models, or handcrafted features that inadequately capture temporal purchasing behaviour. The proposed framework addresses this limitation by integrating comprehensive data preprocessing, RFM-based behavioural feature engineering, and an LSTM network capable of learning long-term customer purchase dependencies. Unlike traditional approaches that perform either segmentation or classification independently, this study combines behavioural feature extraction and deep sequential learning within a unified framework for customer classification. The proposed model achieves high predictive ACC while generating actionable customer segments that effectively support personalized marketing and customer retention strategies in e-commerce platforms.

C. Organization of the Paper

This paper is structured as follows for the rest of it. The literature review is included in Section II. The suggested approach is detailed in Section III. The findings of the experiments and a comparison of the outcomes are presented in Section IV. In Section V, the report comes to a close and provides an outline of potential areas for further research.

2. Literature Review

The literature review examines recent DL-based customer segmentation techniques, highlighting key methods, evaluation metrics, challenges, and research gaps in targeted e-commerce marketing.

R. S. Manikandan et al.,(2026) suggest ESCSODL-CSPBA, a method for improving sand kat swarm optimisation using DL to segment customers based on their purchase behaviour analytics. To start, use the RFM-RN model, which stands for Recency, Frequency, Monetary, and Repurchasing Number of Times. Clustering and feature subset selection are also handled by the ESCSO and DBSCAN models. The classification process is completed by employing the CRNN method. The ESCSODL-CSPBA method outperformed the state-of-the-art models on the online retail dataset, achieving an impressive ACC rate of 95.0 per cent [15].

P.-Y. Pai et al., (2026) introduce a new approach to e-commerce consumer value segmentation by integrating RFM esearch, ML, and association rule mining. Behavioural patterns and sustainability participation are assessed using digital transaction data. Proven by a sustainable food business targeting low-income consumers. Combining Random Forest and XGBoost into an ensemble model yields an AUC of 0.9953 and an ACC of 97.19% [16].

M. Xu (2025) emphasises the use of K-means clustering as a powerful tool for e-commerce PRE marketing. Businesses may now run more relevant targeted advertising thanks to the suggested model's usage of K-means clustering to efficiently split clients into various groups. Results from experiments demonstrate a considerable improvement over previous methods, with a 98.2% ACC rate, 97.7% F1, and better segmentation quality [17].

K. Kim et al., (2025) offer the RFMVDA model, which stands for "Recency, Frequency, Monetary, Visits, Durations, Actions." Data on online shoppers, their sessions, and their actions may be captured using this approach. A DNN was built to forecast consumer behaviour-based classifications using the RFMVDA model for behaviour-based segmentation and classification. Consequently, the suggested model achieved a segmentation prediction ACC of 92.98% for online shoppers [18].

S. R. Ahmed et al., (2024) suggest an innovative architecture that improves e-commerce CRM by utilising DL approaches. CRM activities including churn prediction, recommendation systems, and ANN creation and assessment are at the heart of the strategy. In terms of enhancing predicting ACC in CRM activities and capturing complex client behaviours, the results show that DL is effective. At 90%, the maximum ACC was reached [19].

M. Al Rafi (2022) proposes a data-driven consumer segmentation approach that boosts digital marketing and targeted advertising in today's online marketplaces. K-Means (k=5) outperformed DBSCAN (k=2) in terms of segmentation quality, as measured by the Silhouette Score (0.71) and the Davies-Bouldin Index (0.42). With an F1-score of 0.87 and an AUC of 0.93, a GBM classifier was trained to predict the likelihood of ad clicks in clusters, which is used to measure the effect of marketing [20].

Existing studies have shown that ML and DL techniques can effectively improve customer segmentation and targeted marketing. However, many existing methods rely on traditional clustering algorithms, handcrafted features, or conventional neural networks that are unable to fully capture the sequential and long-term nature of customer purchasing behaviour. They also suffer from high computational complexity, limited generalization, and validation on restricted datasets. As summarized in Table I, these research gaps motivate the development of an efficient LSTM-based DL framework to better learn customer transaction patterns, enhance segmentation ACC, and support more effective targeted marketing in e-commerce platforms.

TABLE Comparative Analysis of Recent Studies on Customer Segmentation Approaches for Targeted Marketing

Authors	Approach	Dataset	Key Findings	Limitations	Recommendation
R. S. Manikandan et al. (2026)	ESCSO + DBSCAN + CRNN	Online Retail Dataset	Achieved 95.90% accuracy for customer segmentation and purchase behavior classification.	High computational complexity and validated only on a single retail dataset.	Develop lightweight deep learning models and evaluate on larger, diverse e-commerce datasets.
P.-Y. Pai et al. (2026)	RFM + Association Rule Mining + RF + XGBoost	Sustainable food e-commerce dataset	Achieved 97.19% accuracy with AUC = 0.9953 for customer value segmentation.	Limited to sustainable retail and does not effectively capture temporal purchasing behaviour.	Integrate deep learning models for sequential customer behaviour analysis across multiple domains.
M. Xu (2025)	K-Means Clustering	E-commerce customer dataset	Obtained 98.2% accuracy and 97.7% F1-score for customer segmentation.	Requires predefined clusters and struggles with complex nonlinear customer patterns.	Employ deep clustering methods with automatic feature learning.
K. Kim et al. (2025)	RFMVDA + Deep Neural Network	E-commerce customer behaviour dataset	Achieved 92.98% classification accuracy using behavioural features.	Relies on handcrafted features and extensive feature engineering.	Uses hybrid deep learning architectures for automatic feature extraction and representation learning.

S. R. Ahmed et al. (2024)	ANN + DNN	Customer transaction dataset	Achieved 90% accuracy in CRM tasks, including segmentation and churn prediction.	Focused on CRM rather than dedicated customer segmentation, with relatively lower predictive performance.	Develop specialized deep learning models for customer segmentation and targeted marketing.
M. Al Rafi (2022)	PCA + K-Means + DBSCAN + GBM	Customer behavioural dataset	Achieved Silhouette = 0.71, F1 = 0.87, and AUC = 0.93 for segmentation and prediction.	Conventional methods require manual feature engineering and have limited capability for complex pattern learning.	Design end-to-end deep learning frameworks for integrated segmentation and purchase prediction.

3. Methodology

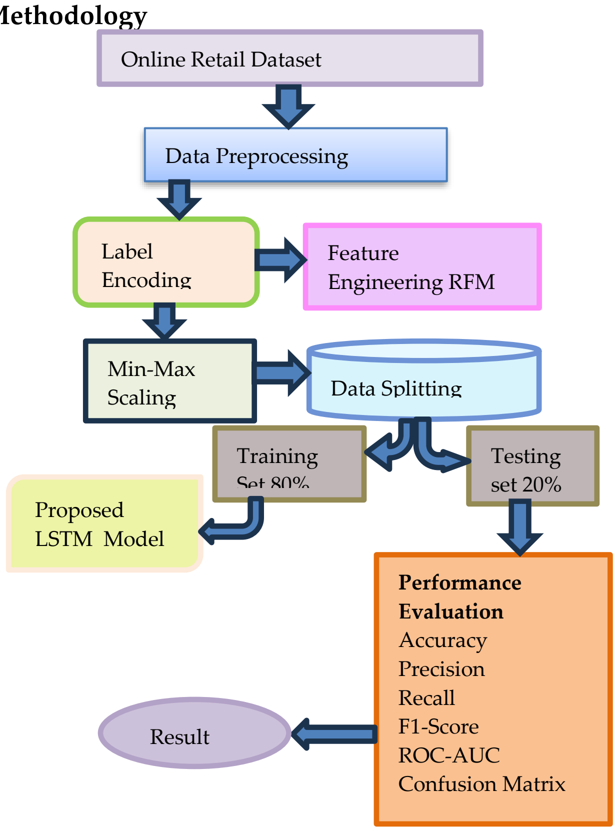

Fig. 1. Proposed DL based Workflow for Targeted Marketing in E-Commerce Platforms

The procedure for consumer segmentation and purchase behaviour prediction using the Online Retail dataset is comprehensively illustrated in Figure 1 by the proposed framework. After gathering transaction data, it is processed to identify incorrect transactions, standardise formats, manage missing values, and delete duplicate entries. Next, label encoding is used to transform categorical characteristics into numerical forms that are more suited for ML models. This phase follows preprocessing. Then, in order to develop useful characteristics based on consumer behaviour, RFM analysis is employed. Then, to make the feature values more consistent and boost the model's convergence, Min-Max scaling is used. As part of the model building and validation process, the processed dataset is split into two subsets: training and testing. The training subset contains 80% of the data while the testing subset contains 20%. Both the training and testing datasets are used to train and assess the proposed LSTM model. At last, the suggested customer segmentation methodology is evaluated for its usefulness using ACC, PRE, REC, F1, ROC-AUC, and the Confusion Matrix.

A. Data Collection

The UCI ML Repository's Online Retail dataset (Dataset ID: 352) [21] is used in this work. It contains 541,909 transaction records obtained from a UK-based online retailer between 2010 and 2011. Here are the eight attributes that make up the dataset: InvoiceNo, StockCode, Description, Quantity, InvoiceDate, UnitPrice, CustomerID, and Country. Used as a standard for consumer segmentation, analysis of purchase behaviour, predictive modelling, and targeted marketing, its extensive transaction history offers important insights into customers' purchasing behaviour.

B. Data Preprocessing

Data preprocessing is an essential procedure that enhances the quality of data by eliminating inconsistencies, errors, and irrelevant records prior to the development of the model. For trustworthy study of consumer behaviour and predictive modelling, it guarantees that the dataset is accurate, consistent, and of high quality.

- Remove duplicate transaction records to eliminate redundant information.
- Customer identity is crucial for customer-level analysis, thus discard entries without it.
- Do not include cancelled invoices with numbers starting with 'C' as they do not reflect finalised purchases.
- Exclude transactions with negative or zero Quantity and UnitPrice values, as they represent invalid

purchase records.

- Convert the InvoiceDate attribute into a standard datetime format to enable temporal feature extraction.

C. Label Encoding

After preprocessing, categorical attributes are converted to numerical representations using label encoding. ML and DL algorithms require numerical input data; therefore, categorical features such as Country or other categorical variables are encoded into unique integer values while preserving their distinct categories. This transformation enables the model to efficiently process categorical information without introducing inconsistencies during model training.

D. Feature Engineering Using RFM Analysis

The RFM (Recency, Frequency, Monetary) model is a core feature-engineering method used to successfully capture client purchasing behaviour. Since RFM is easy to understand and implement, not to mention it effectively measures user engagement and value, it finds extensive usage in customer analytics. Based on transactional data, three important behavioural traits are calculated for each client ci:

Recency: evaluates the time since a consumer last bought anything. It is calculated using Equation (1), and it is the amount of time that has passed from the last transaction to the end of the observation period T:

$$R_i = T - max_j\left(t_{ij}\right) \qquad (1)$$

A customer's frequency is defined as the sum of all of their transactions; Equation (2) shows the frequency:

$$F_i = \sum_j 1 \quad (2)$$

The monetary value represents the sum paid by the consumer and is computed using Equation (3):

$$M_i = \sum_j q_{\{ij\}} \times \mathrm{p}_{\{ij\}} \qquad (3)$$

These three behavioural variables effectively summarise customer purchasing habits and provide a compact yet informative representation of customer value. Before model development, the RFM features are normalized to ensure balanced learning across all variables.

E. Min-Max Feature Scaling

Min-Max Scaling is used to normalise the engineering numerical characteristics such that all input variables are in the range of 0 to 1. Since various consumer behavioural features employ different numerical scales, normalisation helps the neural network converge faster by preventing features with bigger values from dominating the learning process. Equation computes the normalised feature value Equation (4):

$$X_{Scaled} = \frac{X - Xmin}{Xmax - Xmin} \qquad (4)$$

The original feature value is represented by X, whereas the minimum and maximum values of the related feature are X_{min} and X_{max},, respectively.

F. Data Splitting

The generalisability of the suggested model is tested by dividing the dataset into a training set and a test set. 80% of the client records are utilised for model training in this study, with the remaining 20% set aside for testing.

G. Proposed LSTM Model

The suggested architecture makes use of a specialised RNN called an LSTM network to efficiently understand sequential patterns and long-term relationships in consumer purchase behaviour. An input layer, a Softmax output layer, a fully linked dense layer, and one or more hidden LSTM layers make up the LSTM architecture. The input layer receives normalized behavioral feature sequences, while each LSTM cell selectively stores, updates, and discards information using the input, forget, and output gates. The learned hidden representations are forwarded to dense layers with the ReLU activation function to enhance feature learning. In the end, the Softmax layer determines the probability distribution across different types of customers and places each one in the class where the likelihood is highest. The LSTM computations are governed by the following equations:

$$f_t = \sigma(W_f[h_{t-1}, x_t] + b_f) \qquad (5)$$

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \qquad (6)$$

$$\tilde{C}_t = \tanh(W_c[h_{t-1}, x_t] + b_c) \qquad (7)$$

$$C_t = f_t \odot C_{t-1} + i_t \odot \tilde{C}_t \qquad (8)$$

$$o_t = \sigma(W_o[h_{t-1}, x_t] + b_o) \qquad (9)$$

$$h_t = o_t \odot \tanh(C_t) \qquad (10)$$

in which x_t is the input vector at time step t, h_{t-1} is the previous hidden state, C_t is the memory cell state, W and b are the trainable weight matrices and bias vectors, $\sigma(\cdot)$ is the sigmoid activation function, $\tanh(\cdot)$ is the hyperbolic tangent function, and $\odot$ is element-wise multiplication. These gating mechanisms enable the LSTM network to preserve important behavioural information while discarding irrelevant patterns, resulting in effective customer behaviour representation and accurate customer segmentation.

H. Performance Evaluation

Several classification measures, including ACC, PRE, REC, F1, ROC-AUC, Confusion Matrix, and the ROC Curve, are used to check how well the suggested framework works. A thorough evaluation of the suggested model's classification capabilities, robustness, and discriminative power is done using these performance metrics. When everything is said and done, the findings show that the suggested LSTM framework is the best for smart e-commerce client segmentation when compared to other ML and DL methods.

1. Accuracy

The ACC of the model is determined by dividing the total number of categories by the ratio of accurate classifications to total classifications, as stated in Equation (11):

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} \quad (11)$$

2. Precision

The ACC of a positive instance prediction is defined as the percentage of true positives out of all positive examples predicted. It is useful for determining if the model can reduce the number of false positives, as shown in Equation (12):

$$Precision = \frac{TP}{TP+FP} \quad (12)$$

3. Recall (Sensitivity or True Positive Rate)

The percentage of true positives that were accurately forecasted is called REC. Equation (13) shows how it aids in evaluating the model's false negative minimising capabilities:

$$Recall = \frac{TP}{TP+FN} \quad (13)$$

4. F1- Score

The F1 is a well-rounded evaluation measure that takes both REC and ACC into account; it is the harmonic mean of the two. The F1 is determined by Equation (14)

$$F1\_Score = \frac{2\,TP}{2\,TP+FP+FN} \quad (14)$$

where TP stands for true positive and TN for true negative. TP is the number of positively anticipated cases and TN is the number of negatively predicted instances. Contrarily, FP stands for false positive and FN for false negative, respectively, where FP is the number of positive cases that were mistakenly anticipated.

5. Result Analysis and Discussion

An Intel Core i7 CPU, 64 GB of RAM, and Windows 10 made up the experimental environment. The development was carried out in Jupyter Notebook, with data preparation, model training, and performance assessment done in Python and its related libraries, such as TensorFlow, Keras, NumPy, Pandas, Matplotlib, and scikit-learn. The suggested LSTM model for customer segmentation was evaluated and the results are shown in Table II. The model's impressive performance in effectively segmenting customers for targeted marketing applications was demonstrated by its 99.77% ACC, 99.39% PRE, 99.93% REC, 99.66% F1, and 99.94% ROC-AUC score.

TABLE Performance Metrics of the Proposed LSTM Model

Performance Metric	LSTM
Accuracy	99.77
Precision	99.39
Recall	99.93
F1-Score	99.66
ROC-AUC Score	99.94

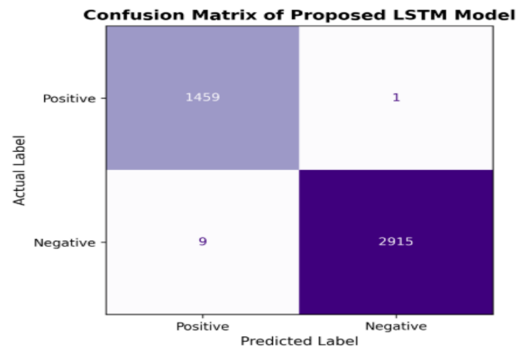

Fig. 2. Confusion Matrix of the LSTM model

Figure 2 shows the suggested LSTM model's classification performance by comparing the actual and projected labels. A total of 1459 positive samples were appropriately categorised as True Positives and 2,911 negative samples as True Negatives by the model. While nine negative samples were erroneously labelled as positive, just one positive sample was erroneously classed as negative (False Negative). It is evident from the small number of examples that were incorrectly identified that the suggested LSTM model accomplishes excellent classification ACC with few prediction mistakes.

Figure 3 shows the suggested LSTM model's learning performance across 50 epochs. The model successfully learns the underlying patterns from the Online Retail dataset, as evidenced by the increasing ACC of both the training and validation sets with each epoch. At around 99.77%, the validation ACC is quite similar to the training ACC, showing strong generalisability.

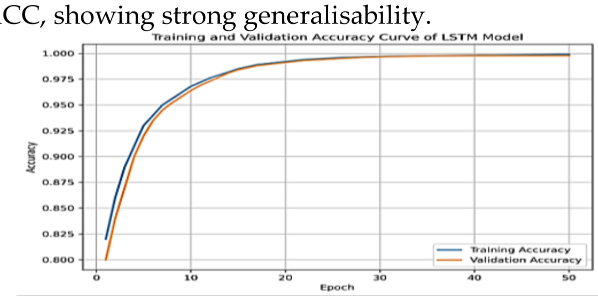

Fig. 3. Training and Validation Accuracy Curve of the Proposed LSTM Model

Figure 4 shows the decrease in model error that occurred during training. That the LSTM model optimises its parameters and improves prediction performance is demonstrated by the steady drop of loss values as the number of epochs grows. It is a sign of good learning and stable model behaviour when the training loss and validation loss converge closely. The suggested model reliably converges, as shown by the constant loss decrease.

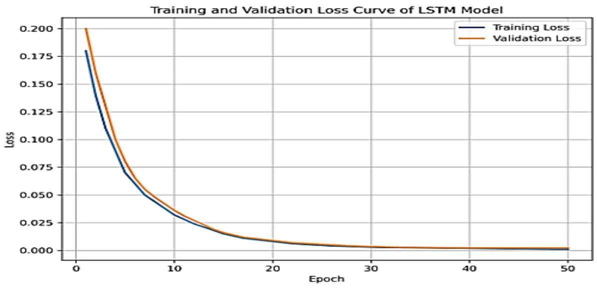

Fig. 4. Training and Validation Loss Curve of the Proposed LSTM Model

Table III shows the results of comparing the suggested LSTM-based consumer segmentation model's performance to that of several previously published ML methods. The findings demonstrate that the suggested model outperforms the competition with a 99.77% ACC rate, 99.39% PRE rate, 99.93% REC rate, 99.64% F1, and 99.94% ROC-AUC. For precise consumer segmentation and focused marketing on e-commerce platforms, the suggested LSTM outperforms LR, XGBoost, and Decision Tree models in terms of predictive power and classification reliability

TABLE III Comparative Performance Analysis of the Proposed Model with Existing Customer Segmentation Approaches

Reference	Dataset	Model	Accuracy	Precision	Recall	F1-Score	ROC AUC
Proposed Model	Online Retail Dataset	LSTM	99.77	99.39	99.93	99.64	99.94
[22]	Real-world E-commerce Dataset	Logistic Regression	72.1	70.5	74	72.2	0.78
[23]	Retail customer Dataset	XGB Classifier	92.40	92.27	92.40	92.28	97.39
[24]	E-Commerce Dataset	Decision tree	0.78	0.79	0.78	0.81	-

5. Conclusion and Future Work

The suggested LSTM-based DL framework is a good way for e-commerce platforms to divide customers into groups and send them targeted ads by using transaction data and behavioural analysis. The framework integrates data preprocessing, label encoding, RFM feature engineering, Min–Max scaling, and LSTM-based sequential learning to accurately classify customers into meaningful behavioural segments. Experimental results demonstrate outstanding predictive performance, achieving 99.77% ACC, 99.39% PRE, 99.93% REC, 99.66% F1, and 99.94% ROC-AUC, confirming its capability to capture complex purchasing patterns and support intelligent marketing decisions. The identified customer segments enable businesses to implement personalized marketing campaigns, improve customer retention, optimize resource allocation, and strengthen customer relationship management. The suggested framework surpasses current consumer segmentation methods in terms of categorisation ACC and resilience, as shown by comparative analysis. A realistic and scalable solution for modern e-commerce analytics is offered by integrating behavioural feature engineering with deep sequential learning. To further enhance model transparency, interpretability, and decision support for personalised marketing applications, future work will seek to validate the framework using larger and more diverse e-commerce datasets, incorporate demographic and clickstream information, investigate transformer-based and hybrid DL architectures, enable real-time customer segmentation, and integrate explainable artificial intelligence techniques.

REFERENCES

1. A. Warrier, "Agentic AI for Natural Language Query Interface in Intelligent Customer Success Management- Conversational Analytics and Automated Reporting," J. Artif. Intell. Cloud Comput., vol. 450, pp. 1–6, Aug. 2025, doi: 10.47363/JAICC/2025(4)450.

2. 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.

3. Y. Patel, "Detection of Multi-Account Abuse in E-Commerce Systems Using Behavioral Biometrics," Am. J. Cogn. Comput. AI Syst., vol. 8, pp. 107–127, 2024.

4. K. Dixit, "Predictive Analytics in Business

Intelligence for Sales Forecasting," Int. J. Adv. Res. Sci. Commun. Technol., vol. 3, no. 1, p. 981, Sep. 2023, doi: 10.48175/IJARSCT-12750G.

5. A. Musunuri, "Leveraging AI and Deep Learning for ECommerce Customer Segmentation," Int. J. Innov. Res. Sci. Eng. Technol., vol. 12, no. 06, pp. 9081–9096, 2023, doi: 10.15680/ijirset.2023.1206165.
6. R. Fu, B. Zheng, J. Wen, and L. Xie, "Research on commodity business value and customer value of e-commerce platforms: Based on consumer psychology and cognition," Front. Psychol., vol. 13, pp. 1–15, Sep. 2022, doi: 10.3389/fpsyg.2022.985537.
7. S. Kukar, "Explainable AI Models for Transparent and Ethical Customer Relationship Management Decision," in 2026 IEEE 2nd International Conference on Secure IoT, Assured and Trusted Computing (SATC), Houston, TX, USA: IEEE, Mar. 2026, pp. 1–7. doi: 10.1109/SATC69565.2026.11542554.
8. A. Joon, S. A. Pahune, S. Mathur, and S. Rongala, "Designing Predictive Analytics Frameworks with ML for ERP-Supported E-Commerce Systems," in 2025 International Conference on Advancements in Smart, Secure and Intelligent Computing (ASSIC), 2025, pp. 1–6. doi: 10.1109/ASSIC64892.2025.11158452.
9. S. U. Rahaman, "Data-Driven Customer Segmentation: Advancing Precision Marketing through Analytics and Machine Learning Techniques," J. Artif. Intell. Mach. Learn. Data Sci., vol. 1, no. 1, pp. 1356–1362, Oct. 2022, doi: 10.51219/JAIMLD/shafeeq-ur-rahaman/309.
10. S. krishna Anumula, "Optimizing Supply Chain Management with AI-Powered Predictive Analytics," J. Int. Commer. Law Technol., vol. 6, no. 1, pp. 244–252, Sep. 2025, doi: 10.61336/Jiclt/25-01-20.
11. D. Patel, "Hybrid Deep Learning Approaches for Personalized Product Recommendation in Online Retail Systems," in 2026 6th International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), FEZ, Morocco: IEEE, 2026, pp. 1–6, May. doi: 10.1109/IRASET68627.2026.11538894.
12. M. Mittal, "Ai In Retail: Transforming The Customer Experience Through Intelligent Automation,"," Int. Res. J. Mod. Eng. Technol. Sci., vol. 7, no. 3, 2025, doi: 10.56726/IRJMETS69338.
13. K. K. Mohammed, "Leadership Practices of Data Engineering for AI and Machine Learning," Int. J. Sci. Res. Eng. Trends, vol. 12, no. 1, 2026, doi: 10.5281/zenodo.18677279.
14. S. Kakkar, B. Ojo, D. K. Rai, B. Mani, and A. Kakkar, "Supplier Performance Prediction in Big Data-Driven Smart Supply Chains Using AI-Based Machine Learning Models," in 2026 World Conference on Computational Science and Technology (WcCST), IEEE, Mar. 2026, pp. 123–128. doi: 10.1109/WcCST67302.2026.11496273.
15. R. S. Manikandan, S. S. Kumar, and J. J. Amalraj, "Harnessing enhanced sand cat swarm optimisation with deep learning assisted customer segmentation with purchase behaviour analytics on the retail industry," 2026. doi: 10.1007/s41870-025-03065-z.
16. P.-Y. Pai, S.-W. Lin, and W.-M. Lu, "Integration of association rule mining and RFM analysis with machine learning for e-commerce customer value segmentation: a sustainable retail perspective," Qual. Quant., vol. 60, no. 1, pp. 87–125, 2026, doi: 10.1007/s11135-025-02259-8.
17. M. Xu, "Data-Driven Customer Segmentation in E-Commerce using K-Means Clustering for Personalized Marketing Strategies," in 2025 3rd International Conference on Data Science and Information System (ICDSIS), 2025, pp. 1–8. doi: 10.1109/ICDSIS65355.2025.11071130.
18. K. Kim, M. Jo, I. Ra, and S. Park, "RFMVDA: An Enhanced Deep Learning Approach for Customer Behavior Classification in E-Commerce Environments," IEEE Access, vol. 13, pp. 12527–12541, 2025, doi: 10.1109/ACCESS.2025.3529023.
19. S. R. Ahmed et al., "Deep Learning for Customer Relationship Management in E-commerce," in 2024 International Congress on Human-Computer Interaction, Optimization and Robotic Applications (HORA), 2024, pp. 1–7. doi: 10.1109/HORA61326.2024.10550481.
20. M. Al Rafi, "Intelligent Customer Segmentation: A Data- Driven Framework for Targeted Advertising and Digital Marketing Analytics," Int. J. Res. Publ. Eng. Technol. Manag., vol. 5, no. 5, pp. 7417–7428, 2022, doi: 10.15662/IJRPETM.2022.0505003.
21. V. Ummadisetty, "Online Retail Data Set," 2017.
22. D. Sridevi, "Adaptive Deep Learning Framework for Fair and Scalable E-Commerce Customer Segmentation in Data-Scarce Environments," J. Inf. Syst. Eng. Manag., vol. 10, no. 36s, pp. 1137–1147, 2025, doi: 10.52783/jisem.v10i36s.6715.
23. J. R. N. B. Rajeshwari, Mallikarjuna Reddy Doodipala, K.Kiran Kumar Varma, P. Pattabhi Ram, "Understanding Consumer Behavior in the Retail Sector Using RFM Segmentation and Machine

Learning: An Analysis," Eur. Econ. Lett., vol. 13, no. 3, pp. 412–420, 2024, doi: 10.52783/eel.v14i3.1783.

24. E. Balasundaram, P. Aranganathan, K. S. Annavajjala, R. Sivakumar, M. Arumugam, and A. Vinoth, "A Hybrid Approach for Customer Segmentation and Loyalty Prediction in E-Commerce," Prabandhan Indian J. Manag., vol. 17, no. 10, p. 56, Oct. 2024, doi: 10.17010/pijom/2024/v17i10/173996.