IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, June 2026
Impact Factor: 8.2

Data Mining Techniques for Consumer Behavior Analysis in Digital Fashion Markets

Dr. Dinesh Yadav
Associate Professor, Department of Computer Science and Engineering
Rungta International Skills University, Bhila, CG
executivedirectorcdc@rungta.ac.in

Abstract: *The success of any organization is contingent upon the behavior and incidents of its customers. Moreover, consumer purchasing patterns are essential in ensuring the fast expansion of e-commerce sites, as they will aid in retention and targeted advertising. The study examines behavioral patterns and predicts potential churners using the E-commerce Customer Churn Prediction dataset. Some models, such as the Logistic Regression (LR) and Generative Adversarial Network (GAN), Deep Neural Network (DNN), and the proposed CNN-LSTM model, are applied and tested with the help of some of the main performance indicators, such as accuracy, precision, recall, and F1-score. The dataset is preprocessed through missing value are fixed, one-hot encoding is done, the data are normalized, and the balancing of the classes is performed with the help of SMOTE in order to increase the reliability of the models. The experimental outcomes indicate that the CNN-LSTM model is better than other models, with an accuracy of 98.56, a precision of 98.25, a recall of 99.87, and the F1score of 99.56, which suggests that it is highly effective in impressing the complicated patterns of consumer behavior. The outcomes reveal the usefulness of DL methods in projecting customer churn and aiding data-driven decision-making in e-commerce websites of digital fashion.*

Keywords: Consumer Behavior Analysis, Data Mining Techniques, Machine Learning, CNN-LSTM Model, Digital Fashion Markets, E-commerce customer churn prediction dataset

I. INTRODUCTION

The analysis of consumer behaviour is a vital element of studying how people interact with digital shopping platforms, especially in the fast-developing digital fashion market. Through online shopping, there is a constant generation of vast amounts of customer information by online fashion retailers via browsing, product search, purchase, and feedback [1][2]. This data of behaviour is interpreted to help in determining customer preferences, buying trends and degree of engagement by businesses [3][4][5][6]. With the growing competition in the online fashion industry, a need to learn how consumers make choices arises in a bid to enhance customer satisfaction, customer retention and customized shopping experiences. Most companies followed a product-focused marketing model, where production was the main emphasis to achieve a superior product and reduce manufacturing expenses by focusing on customers based on market share, as per the concept of economies of scale [7][8]. The psychological variables affecting consumer behaviour in the fashion business encompass perceptions of authenticity, emotional attachment to designs, cognitive fallacies in purchase decisions, and economic models balancing value and price. The utility maximization model and other traditional economic models describe the ways in which consumers allocate their expenditure based on the perceived utility, yet they usually do not describe the relationship processes: the social comparison and influence of trends that define fashion selections.

Digital fashion markets are a dynamic space Customer preferences continuously change in line with fashion, social influence and new technologies that are informed by data to learn customer behaviour in digital fashion markets provides invaluable information on specific marketing strategies, product recommendations and general usability of the

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, June 2026
Impact Factor: 8.2

platform Analytical approaches to be applied in order to extract meaningful information out of mass of behavioural data generated by online fashion platforms [9][10][11]. Data mining tools are useful in the analysis of complicated consumer data and in discerning obscure behavioral trends in online fashion markets. Classification, clustering, and association rule mining are some of the techniques that allow a researcher and a company to identify the relationships between customer attributes, purchasing behaviour, and product categories. Data mining is the process by which tens of fields in massive relational databases are identified as having correlations or patterns. It focuses on the process of finding and determining rules that characterize certain and successive patterns in the ever-growing amount of data in contemporary times, more automatic and effective mining techniques are needed [12][13][14]. The contemporary computational methods enable businesses to manage vast amounts of data and identify both behavioral and transactional relationships in digital fashion platforms[15]. The paper's framework is as follows: Section II offers a literature overview on customer behaviour analysis in digital fashion market; Section III delves into the dataset and methods; Section IV presents the results and analysis of the experiments, and Section V concludes with the most relevant discoveries and suggestions for further research.

A. Data Mining Techniques for Customer Behavior Analysis

Data mining is the practice of finding valuable information inside massive databases by analyzing them for patterns, correlations, and trends [16]. Data mining methods make it easier to comprehend how customers engage with goods, services, and brands when they are used in customer behaviour analysis. The study examines some of the most important data mining methods for analyzing consumer behaviour below:

Cluster Analysis

Cluster analysis is a methodology used to group objects into clusters to assess the degree of similarity among objects within a cluster and between clusters. Clustering is a technique that helps in identifying customer segments that share similar behaviors, preferences or attributes when analyzing customer behaviour.

Association Rule Learning

Association rule learning is an algorithm to find interesting relationships between variables in large data sets. It is commonly used by businesses in market basket analysis, which is the process of identifying the products that are often bought Intrigate[17].

Predictive Analytics

The aim of predictive analytics is to evaluate the customer lifetime value (CLV), predict the sales trends, or predict future events or behaviors based on current and historical data through statistical algorithms and ML processes.

Decision Trees

DT is a form of Supervised Learning algorithm that is employed to categorize data according to input variables. The customer behaviour analysis can be conducted using decision trees to predict consumer behaviors and understand the decision-making process.

Sequential Pattern Mining

Sequential pattern mining is used to detect patterns or sequences of events that appear often in a given order in transactional databases[18].

Text Mining and Sentiment Analysis

The goal of Text Mining and Sentiment Analysis is to draw conclusions from large volumes of unstructured textual data, such as reviews, social media posts, and customer service interactions.

Collaborative Filtering

Collaborative filtering is a method that analyzes the similarities between users' preferences or behaviors in order to generate personalized recommendations.

B. Motivation and Contributions of the Study

Consumer behaviour is a dynamic aspect of e-commerce that requires an understanding and prediction in the modern

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429

Volume 6, Issue 4, June 2026

Impact Factor: 8.2

world to be competitive. Conventional approaches find it difficult to capture the complex and dynamically evolving consumer patterns. The research is driven by the necessity to combine predictive analytics with machine learning to successfully analyze customer behaviour data, improve decision-making and facilitate targeted marketing. The goal of such an approach is to enhance customer interaction, streamline business plans and ensure tailored shopping experiences. The study makes the following contributions:

- The use of the E-commerce Customer Churn Prediction dataset to study consumer behaviour trends and predict possible churn in online fashion e-commerce websites.
- Data pre-processing techniques, including handling missing values, one-hot encoding, data normalization, and SMOTE-based data balancing, to improve dataset quality and model reliability.
- The use of the E-commerce Customer Churn Prediction dataset to study consumer behaviour trends and predict possible churn in online fashion e-commerce websites.
- A comparison of performance between the conventional ML models and the new DL model to assess predictive performance in the determination of the churn behaviour.
- A thorough assessment of the models based on the standard performance indicators, including accuracy, precision, recall, F1-score, and ROC-AUC, which guarantee sound and usable prediction of customer churn in online fashion markets.

C. Justification and Novelty

This research aims to predict customer behaviour in e-commerce, with a unique convolutional learning of features and the bidirectional LSTM of sequence data. The ability of this method to address the limitations of the traditional models by modeling long-term relationships as well as local patterns is a justification. The structure is able to deal effectively with data imbalance and feature scaling in combination with pre-processing methods like SMOTE and min-max normalization. Consequently, it provides predictions that are more precise, generalizable and interpretable, which can be used to develop informed strategies for customer retention.

II. LITERATURE REVIEW

This section reviews existing studies on consumer behaviour analysis in digital fashion e-commerce platforms using ML and DL techniques. A comparative summary of datasets, methods, and results is presented in Table I.

Jeong (2026) provides a way for consumers to use fashion curation sites. Using SEM to integrate task-technology with important structures. Using fashion curation sites, a week-long poll of 300 Korean consumers was used to gather data. The results demonstrate that technological parameters have a considerable effect on task-technology fit, effort, social influence, and enabling factors in influencing behavioral intention [19].

Alarfaj et al. (2026) study uses the SDS to build a scalable and efficient IDS for 5G ML techniques to understand Saudi retail customer data. With a 92% silhouette score, clustering techniques used actual data from Tamimi Markets to classify items into budget, midrange, and luxury categories. According to predictive study, XGBoost outperformed traditional regression by 14% and simple neural networks by 9%. These results surpass the conventional retail analytics approaches' 80% accuracy [20].

Dinh and Vu (2025) CNNs such as ResNet50 efficiently extract visual features from fashion images; these features drive consumer behaviour. There are a lot of social, economic, and psychological aspects that go into people's fashion buying decisions, but an integrated ResNet50-GCN model can assist with that by combining graph-based relational modeling with spatial Feature Extraction. Assessed on the DeepFashion dataset, the method produces results in terms of categorization accuracy, providing valuable information for strategic marketing, trend forecasting, sustainable fashion practices, and economic decision-making within the fashion retail industry [21].

Joon et al. (2025) suggests a a predictive analytics framework for ERP-supported e-commerce systems using ML and DL models. Leveraging the Fashion-MNIST dataset (60,000 training and 10,000 testing samples), the dataset underwent pre-processing, including noise removal, logarithmic transformations, normalization, and feature extraction with CNN. Dimensionality reduction via t-SNE aided data visualization. A CNN-based classification model outperformed other models, achieving 92.43% training and 90.42% testing accuracy, compared to Gradient Boosted Decision Trees (76%) the potential of ML-driven predictive analytics in enhancing ERP-based e-commerce systems [22].

Wang (2025) consumer behaviour is too complex and too diverse, it actually puts higher requirements on traditional data analysis methods identify different consumer groups and how to tap into users' potential needs for consumption upgrades. The DBSCAN clustering algorithm to analyze the data of large-scale users on e-commerce platforms. By clustering user consumption behaviors, some different types of consumer groups, some potential trends in consumption, DBSCAN has been used to deal with scenarios with a lot of noisy data and sparse data distribution, compared to the traditional K-means algorithm in terms of clustering accuracy and the level of detail in group division [23].

Liu et al. (2024) make use of 33,040,175 data from the buys file of actual e-commerce consumers, using a mix of ML and DL models, such as RF, GBDT, XGBOOST, RNN, and LSTM, to predict the buying behaviour of clients with an accuracy of 70–75%. The ML model achieves the most accurate prediction. The most accurate forecast of a customer's propensity to buy a product in the DL model is shown by the 50-layer LSTM model [24].

Lastly, Sharma and Waoo (2023) address this. E-commerce is one of the fastest-expanding industries in the Internet era. People are eager to buy products from online shops such as Amazon, eBay, and Flipkart. Customers and sellers may use online sites to make decisions about marketing tactics and product and service improvements. Consumer behaviour models are often built on ML and data mining of customer data. Predicting consumer behaviour and constructing customer behaviour models requires the appropriate strategy and approach [25].

TABLE 1: SUMMARY OF THE RELATED WORK ON CUSTOMER BEHAVIOR ANALYSIS IN DIGITAL FASHION USING MACHINE LEARNING

Author (Year)	**Dataset**	**Methodology**	**Results**	**Limitation**	**Future Work**
Jeong (2026)	Survey data of 300 Korean consumers using fashion curation platforms	Structural Equation Modeling (SEM) integrating task–technology fit framework	Technology characteristics significantly influence behavioral intention through effort expectancy	Limited sample size and geographical focus on Korean consumers	Expand research to larger and diverse global consumer datasets
Alarfaj et al. (2026)	Real retail transaction data from Tamimi Markets	Machine learning-based consumer analytics with clustering and XGBoost prediction	Achieved 92% silhouette score for product clustering; XGBoost showed 14% lower error	Focused on retail product segmentation rather than deep behavioral modeling	Integrate deep learning models for improved predictive analytics
Dinh and Vu (2025)	DeepFashion dataset	Hybrid ResNet50 + Graph Convolutional Network (GCN) for fashion preference analysis	Improved classification accuracy by combining visual feature extraction with relational modeling	Computational complexity due to deep model architecture	Explore lightweight models for real-time recommendation systems
Joon et al. (2025)	Fashion-MNIST dataset (60k training, 10k testing	CNN-based predictive analytics with preprocessing, normalization, and t-	CNN achieved 92.43% training accuracy and 90.42% testing accuracy,	Dataset limited to image classification rather than real	Integrate transactional and behavioral data for better e-commerce

	samples)	SNE dimensionality reduction	outperforming gradient boosted trees	consumer behavior	analytics
Wang (2025)	Large-scale e-commerce user consumption dataset	DBSCAN clustering for identifying consumer behavior patterns	DBSCAN outperformed K-means in handling noisy and sparse data distributions	Lack of predictive modeling for future purchasing behavior	Combine clustering with predictive models for better consumer insight
Liu et al. (2024)	Real e-commerce dataset with 33,040,175 purchase records	ML models (RF, GBDT, XGBoost) and DL models (RNN, LSTM) for purchase prediction	ML and DL models achieved 72–75% accuracy; LSTM showed best performance in purchase prediction	Moderate prediction accuracy due to complex consumer behavior patterns	Incorporate advanced hybrid models to improve prediction performance
Sharma and Waoo (2023)	E-commerce customer datasets from online platforms	Machine learning and data mining techniques for consumer behavior modeling	Demonstrated the importance of ML-based approaches in predicting customer purchasing behavior	Limited focus on advanced deep learning techniques	Implement deep learning models for improved consumer behavior prediction

III. METHODOLOGY

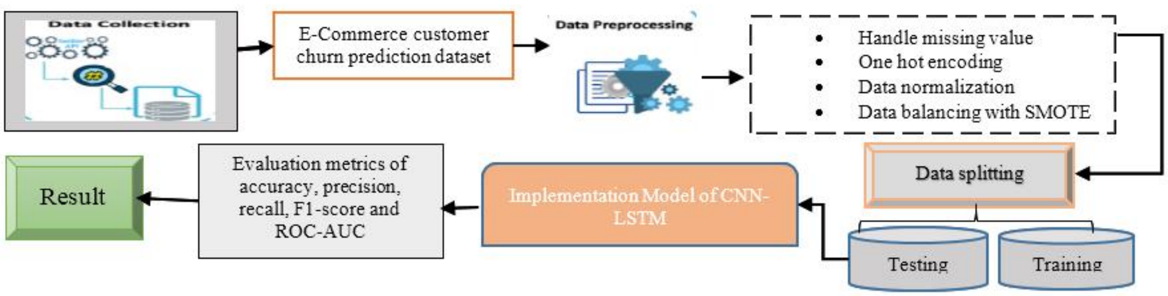

Fig. 1. Flowchart for Customer behavior analysis using machine learning

The e-commerce Customer Churn prediction dataset is utilized to construct a consumer behaviour research framework intended for use in online fashion marketplaces (Figure 1)[26]. It starts with the acquisition of the dataset, then goes through pre-processing, which enhances the quality of data by addressing missing values, encoding categorical features and normalization. To manage imbalance between churn and non-churn customer records[27], the SMOTE technique is applied to generate balanced training samples. The dataset is further refined into training and testing samples to aid in the development of reliable models. Subsequently, the CNNLSTM model is used to acquire complex behavioral patterns and time dependence within customer activity data. Lastly, the system is used to evaluate predictive performance based on accuracy, precision, recall, F1score as well as ROC-AUC, ensuring effective identification of potential churn and supporting better consumer behaviour insights in online fashion marketplaces

Each step of the proposed flowchart of methodology is briefly explained below:

A. Data Collection

The dataset that has been used in this study is the E-commerce Customer Churn Prediction dataset retrieved from Kaggle[28][29]. The processed data has 5,597 records. These records consist of attributes and habits of customers, including their tenure, preferred login device, city tier, and satisfaction score and feature description is listed in Table II. A binary indicator of a customer's churn status (1) or continued activity (0) is the desired variable.

TABLE 2: FEATURE DESCRIPTION OF THE E-COMMERCE CUSTOMER CHURN PREDICTION DATASET

Feature	Description
CustomerID	Unique identifier of each customer
CustomerChurn	Indicates whether the customer leaves the platform (1 = churn, 0 = active)
Tenure	Duration of customer association with the platform
HourSpendOnApp	Average time spent by the customer on the application
OrderCount	Total number of orders placed by the customer
CouponUsed	Number of coupons used during purchases
SatisfactionScore	Customer satisfaction rating
DaySinceLastOrder	Number of days since the last order
CashbackAmount	Cashback received from purchases

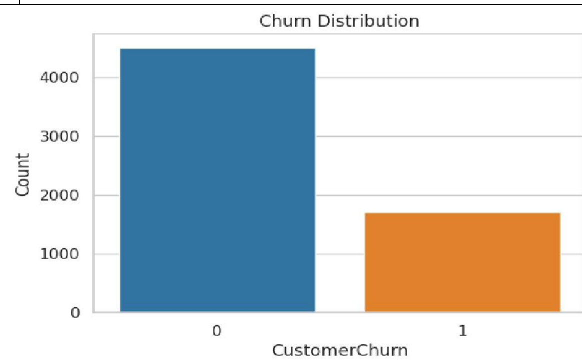

Fig. 2. Count Plot for Class Distribution

The distribution of the class Customer Churn in the e-commerce churn prediction dataset is displayed in Figure 2. The data indicates that there is an evident imbalance in the classes with about 4,500 customers being non-churners (0) and 1,700 customers being churners (1). This asymmetry the usefulness of data balancing methods prior to model training to minimize majority bias and enhance the capacity of the model to effectively.

The Correlation Heatmap of numerical features is shown in Figure 3. The heatmap indicates the correlation between the variables, including Tenure, Order Count, Coupon Used, Day Since Last Order, and Cashback Amount. The majority of features are correlated with each other at the low and middle levels, which means that each attribute is unique to analyze customer behaviour.

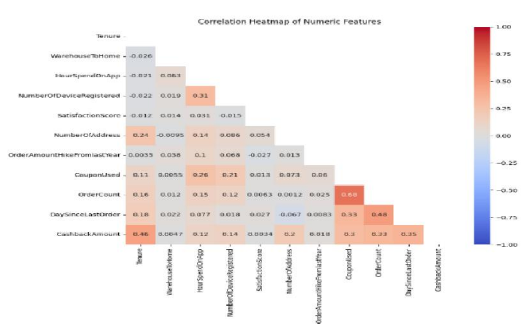

Fig. 3. Correlation Heatmap for Numerical Values

B. Data Preprocessing

The Data Pre-processing Module guarantees the best data preparation by missing data and encoding. Further pre-processing steps are given in next steps.

- **Handling missing values:** This phase is used to discover and deal with missing or null values to make the data complete and accurate. It is vital in preserving credible data, which is pegged on customer data, avoiding bias, enhancing model performance, and allowing the coherent analysis of consumer behavioral patterns.
- **One-Hot Encoding:** One-hot encoding is used to encode non-numeric elements of customer behavior, gender or geography, into a numerical representation that can be utilized in ML models by turning categorical variables into binary vectors of 0s and 1s.

C. Data Normalization

Normalization is a vital pre-processing technique that is applied to bring the numerical features to a standardized range, i.e. between 0 and 1. In this customer behaviour study, a processing technique is applied to bring the numerical features to a standardized range, i.e., between 0 and 1. In this customer behaviour study, min-max normalization is applied using Equation (1)

$$Normalize(X) = \frac{X - X_{min}}{X_{max} - X_{min}} \qquad (1)$$

This method improves convergence speed and predicts accuracy while making sure that all behavioral characteristics.

D. Data Balancing with SMOTE

The outcome of addressing class imbalance, a critical pre-processing step for churn prediction. Using SMOTE, the minority class (churn) are augmented to match the majority class, resulting in an approximately equal distribution of non-churn (0) and churn (1) instances. Balancing the dataset mitigates model bias toward the majority class in Figure 4, which enhances the learning of discriminative features, and improves predictive performance for churn identification

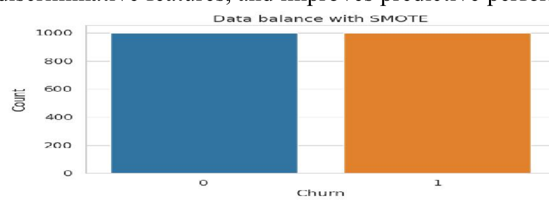

Fig. 4. Data balancing with SMOTE

The outcome of a key pre-processing step in preparing data for churn prediction the distribution of the target variable ‘Churn’ (0 = non-churn, 1 = churn) after applying a balancing technique, resulting in an equal count of 1,000 instances per class. The model's churn forecasts are therefore improved in accuracy and reliability.

E. Data Splitting

The technique of breaking a dataset into smaller pieces so that it may be utilized for testing and training models is known as data splitting. There is an 80:20 split where 80 % of the data shall be allocated to training and 20 % shall be allocated to testing

F. Proposes Models

This section presents the theoretical background of the ML and DL models used for consumer behaviour analysis in digital fashion environment. It briefly explains the classification techniques applied.

KNN

KNN is a supervised ML method that may be applied to both regression and classification. This method is occasionally called the lazy learning strategy because of its minimal learning time demands. KNN determines the unknown value of a new data point by considering feature similarity and assigning the value based on the similarity in the training dataset. The Euclidean distance is a distance measure that is used to measure the distance between affected input and adjacent samples. This involves applying the similarity measures that have been obtained to transform the input data into a single distance measure. This input is thereafter categorized or predicted depending on the mean or median of the labels

DOI: 10.48175/IJARSCT-36573

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 4, June 2026 Impact Factor: 8.2

in KNN in relation to the regression or classification problem. The formula for the distance d in terms of the Euclidean plane between two points v and u is provided by Equation (2)

$$d(v,u) = \sqrt{\sum_{i=1}^{N}(v_i - u_i)^2} \quad (2)$$

Where v_i represents the i^a feature of the input vector and u_i is the ith feature of the training sample.

CNN-LSTM

The CNN component is responsible for extracting local feature representations from customer behavioral data through convolutional operations and pooling layers. These properties are fed into the LSTM layer, which then recognizes sequential connections and temporal dependencies in the input. The model can be trained to detect intricate patterns in customer interactions and purchasing behaviour by combining CNN-based feature extraction with LSTM-based long-term dependency learning. First, the data is fed to Conv1D filters with numerous filters to extract meaningful features of behaviour. A **MaxPooling layer** is applied to reduce dimensionality and computational complexity. The extracted features are then flattened and forwarded to the **LSTM layer,** which captures sequential patterns in customer behaviour data. The output is passed through a Dense layer with ReLU activation once temporal relationships have been learned, followed by a Dropout layer to minimize overfitting. Lastly, Figure 5 illustrates the use of a sigmoid activation function in the output layer

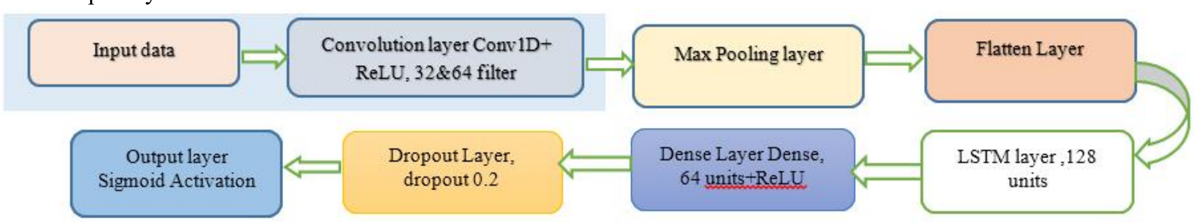

Fig. 5. Architecture of proposed CNN-LSTM model

The CNN-LSTM architecture begins with a 32-and 64-filter convolutional layer activated by ReLU, and continues with a feature-reduction Max Pooling layer. The recovered features are processed using a 128-unit LSTM layer to identify the sequential dependencies in the consumer behavioral data. A Dropout Layer with a rate of 0.2 and a Dense Layer with 64 neurons are incorporated into the model to prevent overfitting. The final Output Layer consists of a Single Neuron with Sigmoid Activation for binary classification. The model is trained using the Adam optimizer with a binary cross-entropy loss function, a Batch Size of 32, and a Learning Rate of 0.0001.

G. Model Evaluation

Accurate evaluation of predictive models is a key to confident e-commerce forecasting. Typically, a confusion matrix is used for this purpose, it provides a transparent picture of model's performance by categorizing predictions as TP, TN, FP, or FN. These results are basis of vital evaluation measures, including acc, prec, rec, F1 and ROC, which are formally specified in Equations (3) through (6) and allow meaningful comparison of various forecasting models.

Accuracy: determines the number of accurate predictions (TP and TN) divided by the total number of predictions.

$$Accuracy = \frac{TP+TN}{TP+FP+TN+FN} \quad (3)$$

Precision: Precision refers to proportion of accurate forecasts to all predictions.

$$Precision = \frac{TP}{TP+FP} \quad (4)$$

Recall: Recall (or Sensitivity) is a percentage measurement of the number of correct identifications of actual positives.

$$Recall = \frac{TP}{TP+FN} \quad (5)$$

F1-score: The harmonic mean of prec and rec is the F1-Score, which balances.

$$F1 - Score = 2 * \frac{(Precision*Recall)}{Precision+Recall} \quad (6)$$

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 4, June 2026 Impact Factor: 8.2

ROC Curve: ROC curve analysis allows calculation of the AUC to assess a model's performance across different threshold values.

IV. RESULT ANALYSIS AND DISCUSSION

The experimental results of the proposed models for customer behavior analysis in digital fashion markets are presented in this section using the e-commerce churn prediction dataset. The tests are carried out on a local Windows 10 machine that has an NVIDIA GTX 1650 GPU, an Intel i5 (9th generation) CPU, and 8 GB of RAM. The implementation is done in Google Colab using tools for data processing, model construction, and visualization including Pandas, NumPy, Scikit-learn, Matplotlib, and Seaborn. The performance comparison shown in Table III demonstrates that both models perform effectively in predicting customer churn. The KNN model has an accuracy of 97.68 percent, whereas the proposed CNN-LSTM has a better performance with an accuracy of 98.56% with a prec of 99.87%, a rec of 99.56%, and the F1score is 98.25%, which shows a strong model that can recognize the trends of customer behavior and forecast attrition in online fashion retailers.

TABLE 3: EXPERIMENT RESULTS OF PROPOSED MODELS FOR CUSTOMER BEHAVIOR IN DIGITAL FASHION MARKET

Measure	KNN	CNN-LSTM
Accuracy	97.68	98.56
Precision	97.63	98.25
Recall	98.21	99.87
F1 Score	97.34	99.56

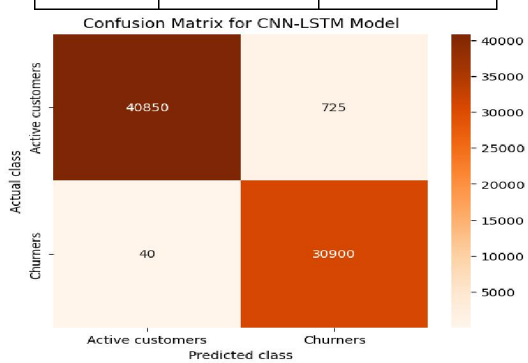

Fig. 6. Confusion Matrix for Proposed Model

The Confusion Matrix of CNN-LSTM model shows strong performance, correctly classifying 40,850 active customers and 30,900 churners, as shown in Figure 6. Only 725 active customers are misclassified as churners, and 40 churners are misclassified as active, indicating high accuracy and very low misclassification.

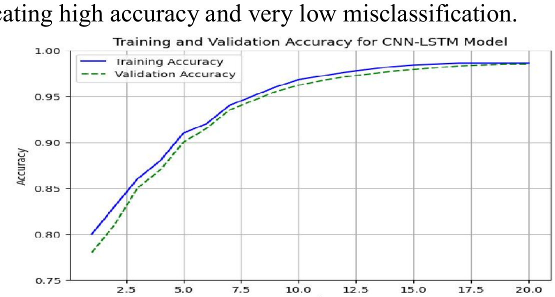

Fig. 7. Training and validation Accuracy curve for CNN-LSTM model

The CNN-LSTM model's performance over 20 epochs of training and validation. Figure 7 indicates that training and validation accuracy steadily increase to nearly 99%, suggesting that the model continues to extract meaningful patterns

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 6, Issue 4, June 2026 Impact Factor: 8.2

from the data.

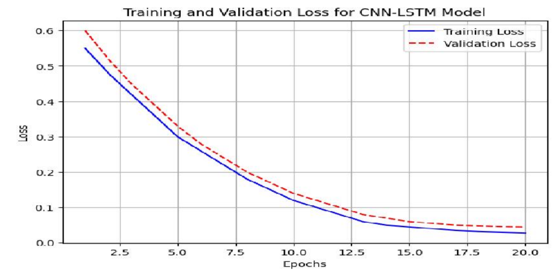

Fig. 8. Training and validation Loss Curve for CNN-LSTM

Figure 8 shows the curve of training and validation loss, which monotonically decreases with the increase in the epochs indicating an improvement in the optimization of the model and convergence stability.

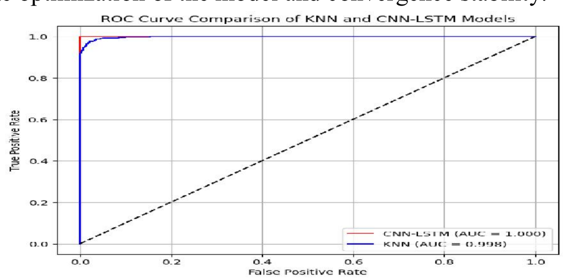

Fig. 9. ROC Curve comparison of KNN and CNN-LSTM

Comparison of ROC curve between the KNN and proposed models used in the analysis of consumer behaviour in digital fashion markets is shown in Figure 9. CNN-LSTM model has AUC of 1.000, indicating that model's categorization performance is flawless, and the KNN model has AUC of 0.998, which is also quite high. which shows a high predictive level in determining customer behaviour patterns.

Comparative Analysis

It compares several ML and DL models for use in digital fashion marketplaces using performance criteria. The performance comparison of several models employing important assessment criteria on the consumer behavior-related e-commerce churn prediction dataset is shown in Table IV. The Logistic Regression (LR) model had 90% accuracy and 91% in classifying the churning behaviour of customers. GAN-based model achieved better predictive performance with 95.79% accuracy, 93.8 precision, 92 recalls, and an F1 score of 92 than baseline model. Accuracy of DNN model are 73.40% with 72.80 precision, 94.60 rec and F1score of 82.30. Conversely, the suggested CNN-LSTM model showed effective results with an accuracy of 98.56, prec of 98.25, recall of 99.87, and F1score of 99.56, which indicates its higher effectiveness in forecasting customer churn, as well as customer behaviour patterns in online fashion markets.

TABLE 4: MODEL PERFORMANCE COMPARISON OF DIFFERENT EVALUATION METRICS ON E-COMMERCE CHURN PREDICTION DATASET

Model	Accuracy	Precision	Recall	F1 Score
LR[30]	90	91	98	94
GAN[26]	95.79	93.8	92	92
DNN[31]	73.40	72.80	94.60	82.30
CNN-LSTM	98.56	98.25	99.87	99.56

The proposed framework enhances consumer behavior analysis integrating ML and DL techniques. The model effectively analyzes customer interaction patterns and buying behavior that enhance accuracy of prediction. The CNN-

DOI: 10.48175/IJARSCT-36573

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 6, Issue 4, June 2026
Impact Factor: 8.2

LSTM model proposed is able to identify the potential churn of customers with 98.56 accuracy, which is reliable. Also, the framework enhances feature extraction by using consumer data to derive temporal and geographical aspects, thereby better understanding user engagement and online fashion platforms' data-driven decisions.

V. CONCLUSION AND FUTURE SCOPE

Online shopping platforms continuously generate data on customer interactions. Consumer behaviour can be critical in helping businesses better interact with customers, making services more effective, and facilitating successful marketing efforts. The digital fashion markets offer various customer browsing, buying trends, product interests and engagement levels, and it can be taken advantage of using analytical methods to extract meaningful insights. The CNN-LSTM model seems to have the greatest performance based on the testing results, with an impressive accuracy of 98.56. The data mining techniques offer formidable means of analyzing these complex datasets and finding concealed patterns that affect consumer decision-making processes of integrating the ML and the DL technologies to analyze the patterns of customer behaviour and help with the predictive analysis in the digital fashion e-commerce context. The presence of various data pre-processing approaches, such as the datasets' missing values, feature encoding, normalization, and balancing, ensures that the data is organized, and it can be used to develop a model with great effectiveness, enhancing the insights into the interaction of consumers with the online fashion platforms and it can be utilized to make decisions based on data to assist online retailers. The research can be improved in future employment by adding real-time behavioral data, more sophisticated hybrid DL models, and explainable AI to enhance consumer behaviour modeling and assist intelligent marketing approaches in online fashion markets.

REFERENCES

[1] C. Demangeot and A. J. Broderick, “Conceptualizing consumer behaviour in online shopping environments,” *Int. J. Retail Distrib. Manag.*, vol. 35, no. 11, pp. 878–894, 2007, doi: 10.1108/09590550710828218.

[2] C. Patel, “A Survey of Data-Driven Customer Segmentation Methods for Targeted Marketing Campaigns,” vol. 3, no. 3, pp. 154–162, 2023, doi: 10.56472/25832646/JETA-V3I7P119.

[3] T. A. Khan *et al.*, “Multi-Source Cyber Intrusion Detection Using Ensemble Machine Learning,” *J. Comput. Sci.*, vol. 21, no. 1, pp. 111–123, Dec. 2024, doi: 10.3844/jcssp. 2025.111.123.

[4] M. H. Moshrefjavadi, H. Rezaie Dolatabadi, M. Nourbakhsh, A. Poursaeedi, and A. Asadollahi, “An Analysis of Factors Affecting on Online Shopping Behavior of Consumers,” *Int. J. Mark. Stud.*, vol. 4, no. 5, 2012, doi: 10.5539/ijms.v4n5p81.

[5] S. Singh, S. A. Pahune, P. Chatterjee, and R. Sura, “Advanced Machine Learning Methods for Churn Prediction and Classification in Telecom Sector,” in *2025 IEEE 6th India Council International Subsections Conference (INDISCON)*, 2025, pp. 1–7. doi: 10.1109/INDISCON66021.2025.11252233.

[6] V. Sikarwar, “AI-Powered Process Mining for Intelligent, Personalized Customer Experience in the Insurance Sector,” *Int. J. Res. Publ. Eng. Technol. Manag.*, vol. 8, no. 4, pp. 12418–12428, 2025, doi: 10.15662/IJRPETM.2025.0804007.

[7] G. V. Bhupathiraju, V. Reddygari, and T. S. Raghavendra, “Data mining for the online retail industry : Customer segmentation and assessment of customers using RFM and k-means,” pp. 1–9, 2022.

[8] T. Shah, “The Role of Customer Data Platforms (CDPs) in Driving Hyper- Personalization in FinTech,” *Int. Res. J. Eng. Sci. Technol. Innov.*, 2025.

[9] M. D. Jain, M. S. Goswami, and M. S. Bhutani, “Consumer Behavior towards Online Shopping: An Empirical Study from Delhi,” *IOSR J. Bus. Manag.*, vol. 16, no. 9, pp. 65–72, 2014, doi: 10.9790/487x-16946572.

[10] N. Prajapati, “The Role of Machine Learning in Big Data Analytics: Tools, Techniques, and Applications,” *ESP J. Eng. Technol. Adv.*, vol. 5, no. 2, 2025, doi: 10.56472/25832646/JETA-V5I2P103.

[11] S. P. Kalava, “Revolutionizing Customer Experience: How CRM Digital Transformation Shapes Business,” *Eur. J. Adv. Eng. Technol.*, p. 4, 2024.

[12] A. Tope-Oke, C. A. Afolalu, and O. Omofade, "A Data Mining-Based Approach to Customer Behaviour in an Electronic Setting," *J. Comput. Commun.*, vol. 07, no. 05, pp. 42–53, 2019, doi: 10.4236/jcc. 2019.75004.

[13] S. H. D. Kolagani and N. S. M. Vuppala, "Human-in-the-Loop and Generative AI Dilemma: A Hybrid Strategy for Effective Customer Service in Enterprise CRM," *Int. J. Bus. Technol. Manag.*, vol. 7, no. 10, Dec. 2025, doi: 10.55057/ijbtm.2025.7.10.18.

[14] T. Shah, "Inbound Vs. Outbound Digital Marketing: A Comparative Analysis of Strategies, Effectiveness, and Business Impact," *Int. J. Res. Appl. Sci. Eng. Technol.*, vol. 13, no. 4, pp. 1–9, Apr. 2025, doi: 10.22214/ijraset.2025.68051.

[15] S. Dodda, N. Kamuni, P. Nutalapati, and J. R. Vummadi, "Intelligent Data Processing for IoT Real-Time Analytics and Predictive Modeling," in *2025 International Conference on Data Science and Its Applications (ICoDSA)*, IEEE, Jul. 2025, pp. 649–654. doi: 10.1109/ICoDSA67155.2025.11157424.

[16] S. K. Pandey, R. Ayyadurai, M. H. Mohammed, K. R. Prasad, S. A. Alisha, and A. Samal, "Data Mining Techniques for Customer Behavior Analysis in E-Commerce," *2024 Int. Conf. Augment. Reality, Intell. Syst. Ind. Autom. ARIIA 2024*, vol. 11, no. 1, pp. 451–457, 2024, doi: 10.1109/ARIIA63345.2024.11051563.

[17] B. Krishnan, S. Perla, S. Maddela, and R. Lingam, "Adaptive Multi-Cloud Infrastructure for CRM Analytics: Real-Time ML and Data Sync with LLMs," in *2025 IEEE 3rd Global Conference on Wireless Computing and Networking (GCWCN)*, IEEE, Nov. 2025, pp. 1–8. doi: 10.1109/GCWCN66157.2025.11448404.

[18] V. Rajendran, D. Besiahgari, S. C. Patil, V. Challagulla, and M. Chandrashekaraiah, "Academic Temporal Pattern Mining for Enhanced Student Performance Prediction: A Hybrid Machine Learning Approach," in *2025 International Conference on Data Science, Agents & Artificial Intelligence (ICDSAAI)*, IEEE, 2025, pp. 1–6, March.

[19] D. Jeong, "How Technology Characteristics and Social Factors Shape Consumer Behavior in Artificial Intelligence-Powered Fashion Curation Platforms," *J. Theor. Appl. Electron. Commer. Res.*, vol. 21, no. 3, 2026, doi: 10.3390/jtaer21030081.

[20] F. K. Alarfaj, M. Badouch, H. U. Khan, and M. Boutaounte, "AIM^2 Framework for Smart Marketing Innovation: AI-Driven Consumer Analytics Using SOR, Neural Networks, and XGBoost in Saudi Retail," Jan. 09, 2026. doi: 10.21203/rs.3.rs-8295953/v1.

[21] H. Dinh and D. A. Vu, "Advances in Consumer Research An Integrated AI Model for Analyzing Consumer Fashion Preferences : Implications for Marketing and Consumer Behavior," *Adv. Consum. Res.*, vol. 2, no. 4, pp. 1–7, 2025.

[22] 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.

[23] W. Wang, "Consumer behavior analysis and consumption upgrade demand mining based on DBSCAN clustering algorithm," in *Proceedings of the 2nd Guangdong-Hong Kong-Macao Greater Bay Area International Conference on Digital Economy and Artificial Intelligence*, ACM, Mar. 2025, pp. 158–163. doi: 10.1145/3745238.3745266.

[24] D. Liu, H. Huang, H. Zhang, X. Luo, and Z. Fan, "Enhancing customer behavior prediction in e-commerce: A comparative analysis of machine learning and deep learning models," *Appl. Comput. Eng.*, vol. 55, pp. 190–204, 2024, doi: 10.54254/2755-2721/55/20241475.

[25] S. Sharma and A. A. Waoo, "Customer Behavior Analysis in E-Commerce using Machine Learning Approach : A Survey," *Int. J. Sci. Res. Comput. Sci. Eng. Inf. Technol.*, vol. 9, no. 2, pp. 163–170, 2023, doi: 10.32628/CSEIT239028.

[26] C. Patel, "Generative AI for Personalized Marketing and Customer Experience in E-Commerce," *Int. J. Emerg. Res. Eng. Technol.*, vol. 7, no. 1, pp. 12–19, 2026, doi: 10.63282/3050-922X.IJERET-V7I1P103.

[27] 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.

[28] H. Ravilla, *Predictive Analytics for Customer Churn in Salesforce Service Cloud*. Springer Nature Switzerland, 2025.

[29] C. Patel, "The Role of Predictive Analytics in Customer Churn Prevention Across Global Markets," *Int. J. Adv. Res. Sci. Commun. Technol.*, vol. 4, no. 2, p. 735, oct. 2024, doi: 10.48175/IJARSCT-19900F.

[30] P. A. Sunarya, U. Rahardja, S. C. Chen, Y. M. Li, and M. Hardini, "Deciphering Digital Social Dynamics: A Comparative Study of Logistic Regression and Random Forest in Predicting E-Commerce Customer Behavior," *J. Appl. Data Sci.*, vol. 5, no. 1, pp. 100–113, 2024, doi: 10.47738/jads.v5i1.155.

[31] R. Akshara, S. K. Reddy, and A. Jain, "Leveraging Deep Learning to Forecast E-Commerce Product Fulfillment," *SSRG Int. J. Electron. Commun. Eng.*, vol. 12, no. 5, pp. 118–125, 2025, doi: 10.14445/23488549/IJECE-V12I5P110.