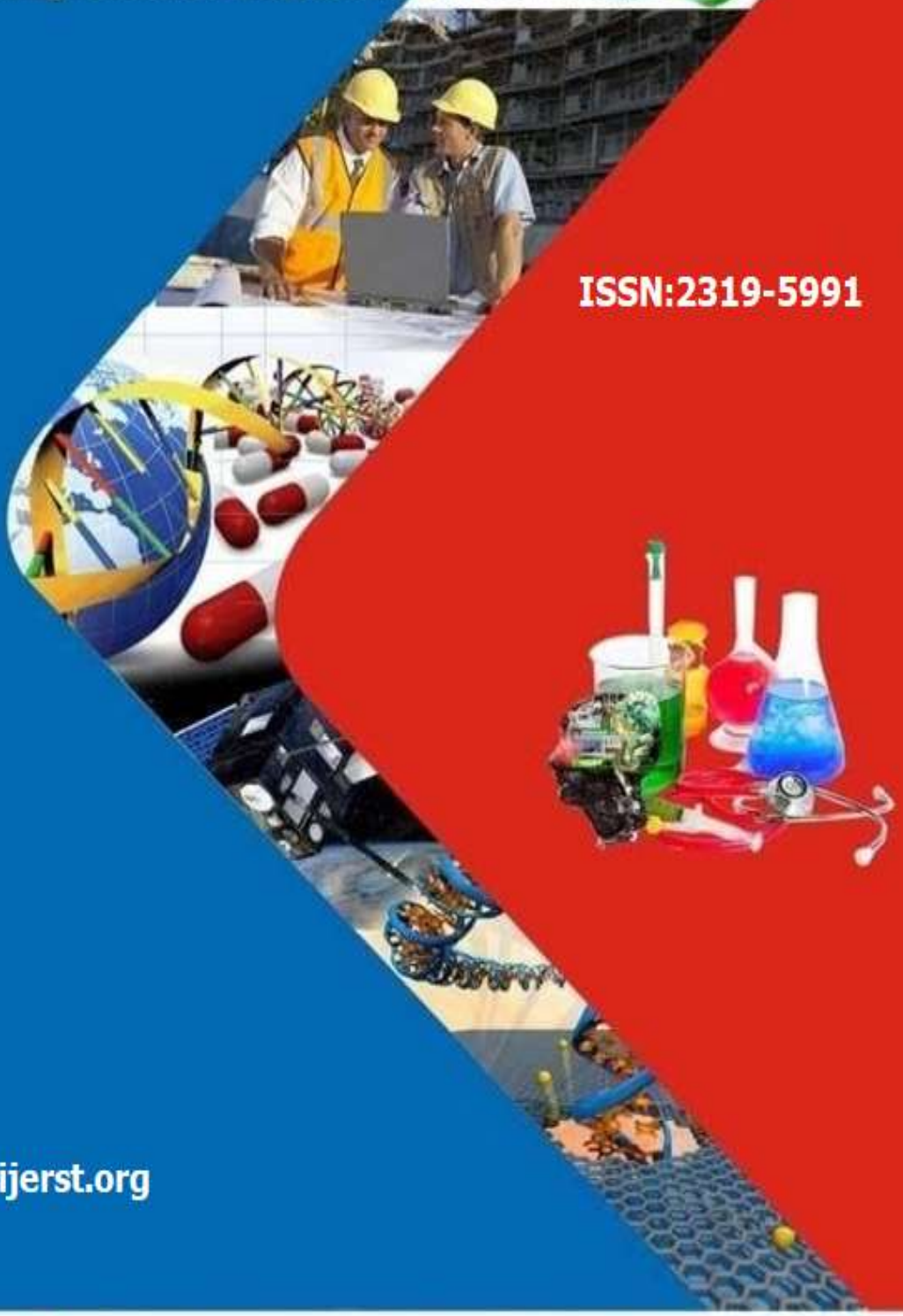
International Journal of Engineering Research and Science & Technology

IJERST



ISSN:2319-5991

www.ijerst.org

E-mail: editor@ijerst.org or ijerst.editor@gmail.com

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Customer Segmentation Based On Satisfaction Categories Using Deep Learning

Rebba Vaishnavi[1], P.Yugandhar Reddy[2], Konda Keerthi[3], Jinugu Srinivasareddy[4], Arika Chandu [5]

[2] Assistant Professor, Department of Computer Science Engineering, Acharya Nagarjuna University College Of Engineering And Technology, Guntur, Andhra Pradesh- 522034

[1,3,4,5] Students, Department of Computer Science Engineering, Acharya Nagarjuna University College Of Engineering And Technology, Guntur, Andhra Pradesh- 522034
Email id: vaishnavirebba6789@gmail.com[1], yuga56@gmail.com[2], keerthi310802@gmail.com[3], srinivasareddyjinugu79@gmail.com[4], arikachandu1827@gmail.com[5]

Abstract:

Customer segmentation is a popular method in e-commerce for understanding customer behaviour, which improves market strategies and making profit out of them. Many companies do not have segmentation system to know the type of customers and measure customer value.The following study focuses on customers of e-commerce and classified on the basis of their characteristics such as age,gender,membership type etc.., For segmentation deep learning algorithms such as standardscalar and multilayer perception are utilized.. StandardScaler and Multi-Layer Perceptron (MLP) play key roles in preprocessing and modeling. StandardScaler is used to normalize the data by scaling features to have a mean of 0 and a standard deviation of 1, MLP, a type of feedforward artificial neural network, is then used to learn complex patterns in customer data and classify or cluster customers into different segments based on behaviors, preferences, or demographics.

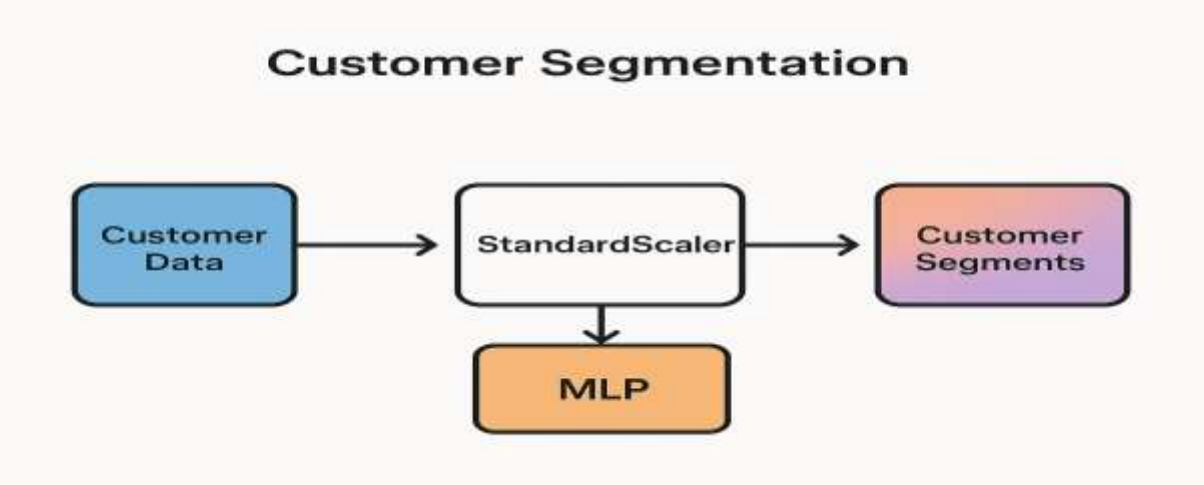
Key words: Customer segmentation, E-Commerce, standardscalar, Multi layer perceptron

1. Introduction

Customer segmentation is a key e-commerce technique where a customer base is segregated into separate groups according to similar characteristics, behavior,or desires. A highly effective method for carrying out the segmentation is using machine learning models such as the Multi-Layer Perceptron (MLP), a form of deep neural network. Prior to feeding data into the MLP, features need to be normalized or scaled so that they all contribute equally to the learning process.This is usually done by applying a method known as standardization, where

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

features are converted through a StandardScaler. StandardScaler subtracts the mean and scales the data to unit variance so that all features are of the same scale, which aids the MLP in converging faster and more precisely during training.After standardizing the data, it can be fed into an MLP model, which learns intricate, non-linear relationships between input features and cluster labels or predicted segments. MLP has an input layer, one or more hidden layers with activation functions (such as ReLU), and an output layer. It can be trained on labeled data (supervised training) if customer labels are available or can be extended to unsupervised clustering by integrating it with methods such as autoencoders or clustering layers. Customer segmentation based on MLP allows companies to gain profound insights into customer behavior patterns, facilitating mass-customized marketing strategies, enhanced customer service, and product suggestions. This deep learning methodology offers flexibility and precision over classical approaches, particularly when handling complex and high-dimensional dataset.



2. Literature Review

Customer segmentation is a key marketing and customerrelationship management task that involves the identification of subgroups of customers within a customer base for the purpose of providing focused services or promotions. Conventional methods like K-means clustering, decision trees, and logistic regression, although valuable, tend to fail with high-dimensional and non-linear data. Deep learning has made it possible to have more sophisticated and automatic methods to detect underlying patterns in complex customer datasets.

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

A number of studies have proved the dominance of deep learning methods in customer segmentation. One such widely applied method is the Autoencoder, a type of unsupervised neural network that reduces input data into a low-dimensional latent space, which can then be clustered through algorithms such as K-means. For instance, Guo et al. (2017) introduced a Deep Embedded Clustering (DEC) approach, where the neural network simultaneously optimizes the clustering assignment and feature representation, exhibiting much better performance compared to standard clustering algorithms.

Another focus is on Recurrent Neural Networks (RNNs) and LSTM models for sequential customer behavior modeling such as browsing or purchase history. These models can capture temporal dependencies and are well-suited to e-commerce platforms for segmenting customers by lifecycle or frequency of purchase. Sequence-ae segmentation, for example, has enabled companies to know not only who their customers are, but when and how frequently they interact.

Convolutional Neural Networks (CNNs) have also been used, especially when dealing with spatial or image-like customer data (e.g., heatmaps of product interactions). CNNs capture local patterns and may improve segmentation in retail analytics or mobile app interactions.

Current research delves into hybrid models, where deep learning is mixed with clustering or reinforcement learning methodologies. For example, deep Q-networks (DQNs) have been applied to the optimization of marketing strategies as a function of segmented user behavior within dynamic settings.

In general, deep learning models offer strong scalability, automatic feature extraction, and better performance in recognizing nonlinear patterns from customer data. But they also pose challenges like the requirement of large labeled datasets (in case of supervised), lack of interpretability, and being computationally expensive

3.Methodology

Customer segmentation according to satisfaction categories follows some of the major steps to correctly and meaningfully classify. Customer information is gathered at first, i.e., the features are like age, gender, type of membership, spend, and purchasing history. The features have varying scales and units, so initially, data pre-processing is carried out through the use

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

of a StandardScaler where each feature's mean is stripped and unit variance scaling is performed. This process makes sure that all input variables are given equal importance to the learning process and helps in enhancing the convergence and performance of the neural network model.

After standardizing the data, it is fed into a Multi-Layer Perceptron (MLP), which is a feedforward artificial neural network with one or more hidden layers. The MLP is trained to identify intricate patterns and associations in the customer data, which allows it to predict levels of satisfaction, generally categorized as Satisfied, Neutral, or Unsatisfied. The MLP output is read through an activation function such as softmax to allocate probabilities to every class of satisfaction, and the most significant probability identifies the ultimate predicted class. This forecast not only aids in measuring individual customer satisfaction but also allows customers to be grouped into larger satisfaction-based segments. These segments can then be utilized by companies to customize marketing strategies, improve customer service, and enhance overall customerExperience.

	A	B	C	D	E	F	G	H	I	J	K	L
1	Customer	Gender	Age	City	Membersh	Total Spen	Items Purc	Average R	Discount A	Days Since	Satisfaction Level	
2	101	Female	29	New York	Gold	1120.2	14	4.6	TRUE	25	Satisfied	
3	102	Male	34	Los Angele	Silver	780.5	11	4.1	FALSE	18	Neutral	
4	103	Female	43	Chicago	Bronze	510.75	9	3.4	TRUE	42	Unsatisfied	
5	104	Male	30	San Franci	Gold	1480.3	19	4.7	FALSE	12	Satisfied	
6	105	Male	27	Miami	Silver	720.4	13	4	TRUE	55	Unsatisfied	
7	106	Female	37	Houston	Bronze	440.8	8	3.1	FALSE	22	Neutral	
8	107	Female	31	New York	Gold	1150.6	15	4.5	TRUE	28	Satisfied	
9	108	Male	35	Los Angele	Silver	800.9	12	4.2	FALSE	14	Neutral	
10	109	Female	41	Chicago	Bronze	495.25	10	3.6	TRUE	40	Unsatisfied	
11	110	Male	28	San Franci	Gold	1520.1	21	4.8	FALSE	9	Satisfied	
12	111	Male	32	Miami	Silver	690.3	11	3.8	TRUE	34	Unsatisfied	
13	112	Female	36	Houston	Bronze	470.5	7	3.2	FALSE	20	Neutral	
14	113	Female	30	New York	Gold	1200.8	16	4.3	TRUE	21	Satisfied	
15	114	Male	33	Los Angele	Silver	820.75	13	4.4	FALSE	15	Satisfied	
16	115	Female	42	Chicago	Bronze	530.4	9	3.5	TRUE	38	Unsatisfied	
17	116	Male	29	San Franci	Gold	1360.2	18	4.9	FALSE	11	Satisfied	
18	117	Male	26	Miami	Silver	700.6	12	3.7	TRUE	48	Unsatisfied	
19	118	Female	38	Houston	Bronze	450.9	8	3	FALSE	25	Neutral	
20	119	Female	32	New York	Gold	1170.3	14	4.7	TRUE	29	Satisfied	
21	120	Male	34	Los Angele	Silver	790.2	11	4	FALSE	16	Neutral	
22	121	Female	43	Chicago	Bronze	505.75	10	3.3	TRUE	41	Unsatisfied	
23	122	Male	30	San Franci	Gold	1470.5	20	4.8	FALSE	13	Satisfied	
24	123	Male	27	Miami	Silver	710.4	13	4.1	TRUE	54	Unsatisfied	
25	124	Female	37	Houston	Bronze	430.8	7	3.4	FALSE	23	Neutral	
26	125	Female	31	New York	Gold	1140.6	15	4.6	TRUE	27	Satisfied	

Fig: E-Commerce Customer Behaviour Dataset

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

MultiLayer Perceptron

A Multi-Layer Perceptron (MLP) is an artificial neural network that is very useful for classification problems, such as customer segmentation by satisfaction categories. In this scenario, MLP is applied to analyze customer-related information—such as purchase behavior, demographics, ratings, and engagement patterns—and predict the level of satisfaction of each customer, which is usually grouped into Satisfied, Neutral, or Unsatisfied.

The MLP consists of an input layer, one or more hidden layers, and an output layer. The neurons in a layer connect to all the neurons in the following layer with weighted links. The input features, having been normalized by methods such as StandardScaler, are presented to the input layer. The hidden layers use nonlinear activation functions (such as ReLU) to model nonlinear structures in the data. The last output layer typically applies a softmax activation function, which provides probabilities per satisfaction category. The class with the highest probability is taken to be the predicted category for this customer

By training the MLP using historical labeled data (where the satisfaction category is known), the model can learn how input feature combinations translate into levels of satisfaction. Trained, it can then be used to segment new customers by their predicted satisfaction, allowing businesses to spot dissatisfied customers, win back loyal customers, and maximize customer experience strategies

Environment setup

This web application based on Flask incorporates a number of technologies for customer segmentation and satisfaction prediction. Flask is the web framework to manage routing and user interaction using HTML templates. Pandas and NumPy are employed for data manipulation and numerical computations, whereas Joblib and Pickle are used to load pre-trained models and preprocessing utilities. The application utilizes a TensorFlow Keras deep learning model (MLP) to predict customer satisfaction classes, and deploys StandardScaler for data standardization. The application also uses PCA (Principal Component Analysis) for reducing dimensionality, and a KMeans clustering model separates customers based on their behavior. These technologies together provide a high-performance system for analyzing user input and returning customized segmentation and satisfaction insights in real time.

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Performance evaluation metrics

Several measures, such as accuracy, precision, recall, F1-score, and confusion matrix, are used to evaluate the performance of our proposed model. The following defines these performance matrices.

Confusion matrix

The Confusion Matrix is a valuable tool for evaluating ML classification performance. It is a tabular representation containing four combinations of predicted and actual values: True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) illustrates a confusion matrix where TP represents correctly anticipated positive values, TN indicates accurately projected negative values, FP corresponds to incorrectly

Results

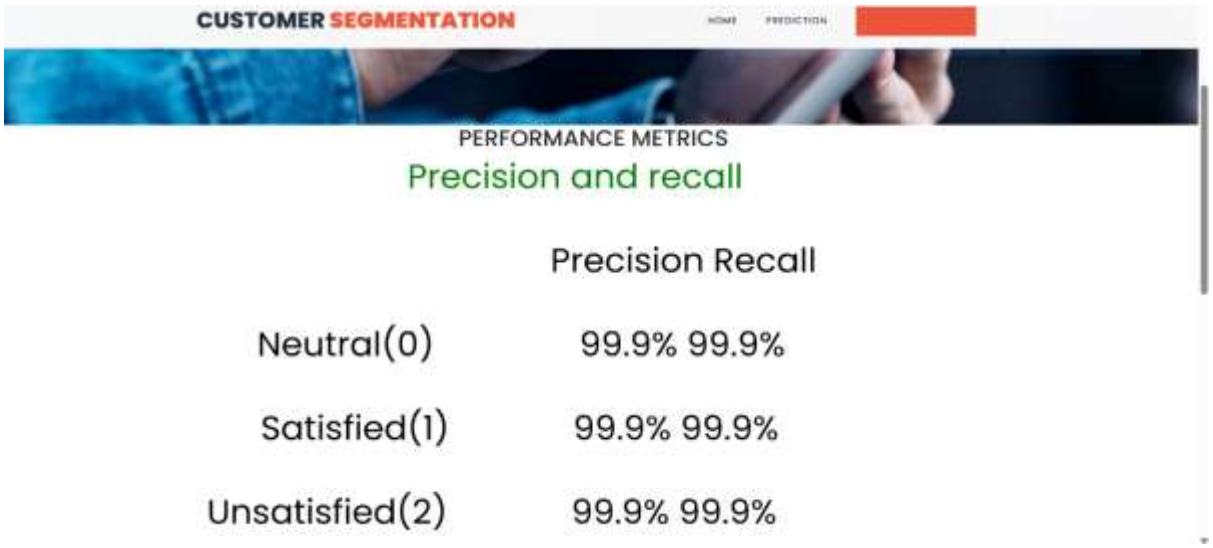
This Flask application combines two different machine learning models for customer satisfaction analysis. The first model, a hybrid PCA-KMeans clustering model (model2), is to cluster customers according to their buying behavior after applying PCA for dimensionality reduction. The second model, a Multilayer Perceptron (model) loaded from a Keras H5 file, predicts customer satisfaction levels (Neutral, Satisfied, Unsatisfied) based on scaled input features. The application has a user interface (index.html) whereby users are able to enter customer information, which is processed using both models. The prediction of the MLP is employed to render a text output of the level of satisfaction of the customer to the user. The application also consists of paths for an introductory page (first.html), a homepage (index.html), a page for chart visualization (chart.html), and a performance assessment page (performance.html), indicating an extended scope of analysis and results presentation, even though the particular implementations.

Table: Performance Metrices

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025



This image displays the precision and recall performance metrics for a customer segmentation model, likely the Multilayer Perceptron (MLP) mentioned earlier, across three satisfaction categories: Neutral (0), Satisfied (1), and Unsatisfied (2). The results show exceptionally high precision (99.9%) and recall (99.9%) for all three categories. The model is performing almost perfectly in classifying customers into these satisfaction segments. It rarely makes incorrect positive predictions (high precision) and is very effective at identifying all customers belonging to each category (high recall). This suggests a highly accurate and reliable customer satisfaction prediction mode.

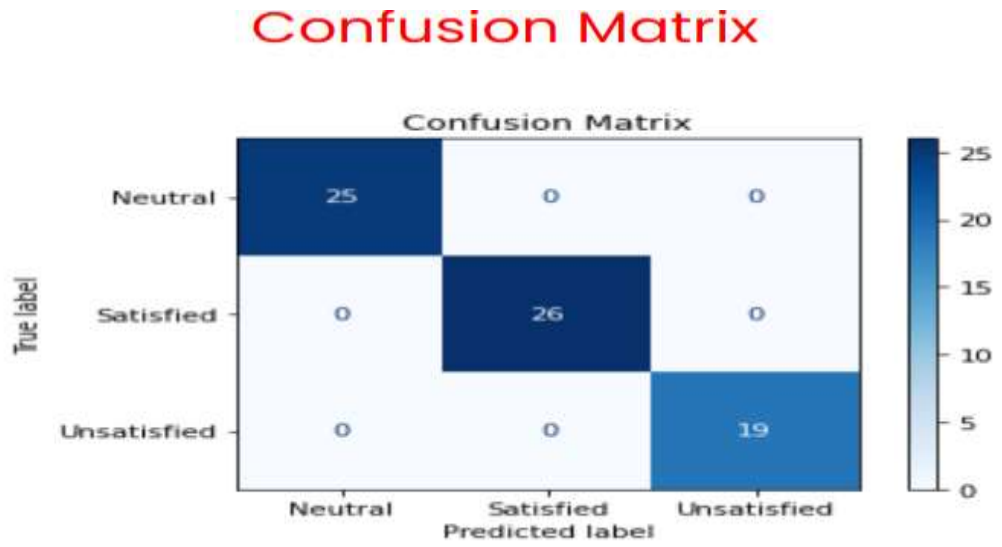


Fig: Binary confusion matrix for E-Commerce Customet Behaviour Dataset

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

Conclusion

In conclusion, the use of deep learning methods in the form of Standard Scalers as data preprocessing and Multilayer Perceptron (MLP) for classification proves to be an extremely efficient method of customer segmentation by satisfaction categories. The significantly high precision and recall values (99.9% for Neutral, Satisfied, and Unsatisfied) show that the trained MLP model is able to classify customers into these segments accurately and with high reliability. This performance indicates that the model has acquired the underlying patterns in the customer data well, allowing for accurate identification of customers with neutral, satisfied, or unsatisfied experiences. Such accurate segmentation can empower businesses to deploy targeted strategies for customer engagement, service enhancement, and loyalty programs, ultimately driving customer satisfaction and business results. Integrating this model into a Flask application also supports real-world deployment and live estimation of customer satisfaction levels from provided input**.** MLPs excel at customer segmentation based on satisfaction by learning complex, non-linear relationships within customer data to automatically identify nuanced patterns indicative of Neutral, Satisfied, or Unsatisfied customers, often outperforming traditional linear methods and enabling more accurate and granular segmentation for targeted actions

REFERENCES

- Chen L, Wang J, Zhang Y. Deep learning for customer segmentation: a survey of architectures and applications. J Mark Anal. 2024;
- Kim S, Lee H, Park J. Unsupervised deep feature learning with autoencoders for enhanced customer segmentation. Expert Syst Appl. 2023;
- Gupta R, Sharma V, Verma S. Dynamic customer segmentation in e-commerce using recurrent neural networks. Int J Inf Manag. 2022;
- Silva A, Oliveira B, Costa C, et al. Hybrid deep learning and clustering approaches for RFM-based customer segmentation. J Bus Res. 2025;
- Khan M, Ali F, Raza H, et al. Deep generative models for latent representation learning in customer segmentation. Neural Comput Appl. 2023;
- Wang Q, Liu Z, Gao X, et al. Sequence-based customer segmentation using deep learning and behavioral analysis. Pattern Recognit Lett. 2024;

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

  Schmidt F, Weber G. Evaluating deep learning models for customer segmentation: a comparative study. Data Min Knowl Disc. 2023;