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 5, Issue 2, November 2025
Impact Factor: 7.67

Ecommerce Customer Segmentation Using Data Science For Targeted Marketing

Prof. Diksha Fulzele, Mr. Vedant Dhote, Ms. Khushboo Patle, Mr. Sujal Dhanwij
Department Computer Science & Engineering,
Tulsiramji Gaikwad Patil College of Engineering and Technology, Nagpur, Maharashtra, India
dikshafulzele73@gmail.com, dhotevedant15@gmail.com
khushipatle01@gmail.com, sujaldhanwij.2022@gmail.com

Abstract: *Customer segmentation is essential in today's e-commerce landscape, helping businesses create targeted marketing efforts and improve customer retention through personalized experiences. This study looks at how machine learning, specifically the K-means clustering algorithm, can be used to identify meaningful customer segments by examining behavioural data such as spending habits and use of discounts from a real-world e-commerce dataset. The process includes data cleaning, creating relevant features, normalizing the data, and using repeated clustering techniques to form distinct groups like "Premium Customers," "Budget Shoppers," and "Value Seekers." Through detailed testing and data visualization, it's shown that focusing marketing efforts on specific segments can significantly boost conversion rates, increase customer lifetime value, and support overall business growth. A comparison with other machine learning and statistical segmentation methods highlights the effectiveness of K-means clustering, making it a reliable tool for large-scale, easy-to-understand market analysis in ever-changing online retail settings. The research also explores how this framework can be adapted for real-time use and integrated with more advanced AI technologies, ensuring its usefulness as e-commerce environments and customer behaviours continue to change..*

Keywords: Customer segmentation, E-commerce, Machine learning, K-means clustering, Behavioural analytics, Targeted marketing, Personalization, Data-driven insights, Customer retention, Conversion optimization, Cluster analysis, Digital marketing

I. INTRODUCTION

Customer segmentation has become a key approach in e-commerce, driven by the need for businesses to keep up with fast-changing consumer demands, new technologies, and fierce competition.

As more digital platforms appear, customers engage with brands through various channels such as websites, mobile apps, social media, email, and online marketplaces, generating huge amounts of data on how they behave. In this fast-moving environment, using a single marketing strategy is no longer enough to form real connections with customers. Companies that don't personalize their approach risk losing out to more flexible competitors.

Segmenting customers means grouping them based on specific factors like their buying habits, how often they shop, how recently they made a purchase, how much they typically spend, their sensitivity to discounts, their demographics, lifestyle choices, and even their personality traits.

E-commerce brands use these groupings to create focused marketing efforts, like specific offers, product suggestions, and tailored messages. For example, loyal customers who spend a lot may appreciate special programs and benefits, while customers who look for deals may respond better to limited-time offers and bundles. By using segmentation, marketers can send personalized messages through various channels like email, ads, text messages, and in-app notifications, making their content more relevant and helping increase customer engagement. Personalization through segmentation leads to higher conversion rates.

Research shows that targeted campaigns can lead to up to 200% more conversions than general marketing campaigns. By matching messages to the needs of each group, e-commerce brands can better understand what motivates customers,

remove barriers to buying, and create positive emotional connections. Moreover, personalized strategies help keep customers coming back, which is much cheaper than gaining new customers. Returning customers not only spend more over time but also recommend the brand to others, increasing brand trust through word-of-mouth and social proof.
Segment-based marketing also helps companies use their resources more wisely.
Rather than spreading their marketing budget evenly, businesses can focus on the most valuable customer groups or those most likely to make a purchase. This focused approach improves the return on ad spend and allows for precise measurement of how well campaigns are performing. Most importantly, segmentation gives businesses useful insights, like discovering why some customers abandon their carts, understanding their loyalty patterns, or spotting new interests, which allows for proactive improvements in websites, product selections, and support systems.
In addition to improving direct marketing results, customer segmentation influences long-term business strategies.
Product development, inventory planning, and channel choices become more informed by data, as brands gain a clearer picture of their customer base and how it evolves. The ability to quickly respond to changes in customer behaviours or profitability helps businesses stay agile and innovative. With the help of machine learning, segmentation techniques now include real-time analysis and predictive modelling, allowing for even more personalized and dynamic marketing efforts.

II. LITERATURE SURVEY

Recent studies in e-commerce customer segmentation show that more and more businesses are using machine learning to deal with the complexity of analysing big and varied customer data.
These studies often point to K-means clustering as one of the most effective and easy-to-understand unsupervised learning methods for grouping customers based on their behavior, transaction history, and personal details. Kumar et al. (2025) compared K-means with hierarchical clustering and found that K-means created more distinct and useful customer groups for targeted marketing in e-commerce, with better results in terms of how well the groups are separated, as shown by higher silhouette scores and clearer visualizations. Their work identified customer types like high searchers with low buying habits, loyal customers, and those with moderate engagement, showing that segmentation can uncover important customer types that are key for business strategies.
In addition to clustering methods, researchers are also looking at multi-criteria decision-making tools that work alongside machine learning.
Barrera et al. (2024) developed a system that combines RFM (Recency, Frequency, Monetary) with methods like AHP, PROMETHEE, and advanced sorting techniques to help prioritize and categorize large customer groups. This approach lets companies include their own strategic goals when determining customer segments, making sure that customer profiles better match business objectives. When compared to pure K-means methods, this hybrid approach performed better, especially in B2B healthcare markets, showing that it can be useful across different industries and customer sizes. Many studies also look into improving the features used for segmentation.
Experts suggest including not just transaction data, but also demographic information, personal preferences, and digital behaviours like browsing patterns, how recently someone purchased, and how often they use discounts. Joung et al. (2023) introduced a machine learning model that uses SHAP values to understand which product features are most important from customer reviews. This helps in creating customer segments based on unmet needs rather than just past purchases. Their method overcomes some of the limitations of traditional approaches by focusing on customer sentiment and specific feature insights, making personalization more effective.
New developments include using reinforcement learning and evolutionary algorithms together with clustering to improve the accuracy and adaptability of segmentation.
Wang et al. (2025) proposed K-means-QLDE, which merges Q-learning with differential evolution to make customer segmentation more precise. By using PCA for dimension reduction and optimizing cluster assignments dynamically, their framework achieved over 95% accuracy on test data, showing how reinforcement learning can make unsupervised models better for digital marketing. However, they did mention some challenges with interpretability, but they believe this method has a lot of potential for dealing with complex and high-dimensional e-commerce data. There's also a growing trend towards using hybrid and ensemble models to overcome the weaknesses of individual methods.

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 5, Issue 2, November 2025 Impact Factor: 7.67

Neural network-based clustering has been tested against traditional algorithms, and it shows promise in creating flexible and nonlinear customer segments that capture complex behaviour patterns. Fazlollahtabar and Ghodsypour (2019) showed that support vector machine (SVM) models can separate customer groups with different purchasing behaviours based on product types, buying frequency, and spending, outperforming traditional methods in both flexibility and precision. These combined approaches are better at handling the messy and ever-changing nature of real-world e-commerce data and consumer habits.

Personalization and adaptive modelling are seen as essential for the future of customer segmentation.

Researchers stress the importance of models that can change and adapt to real-time data and changing customer profiles, ensuring marketing stays relevant in highly competitive e-commerce markets. Arora and Souza (2022) used multi-criteria decision-making with machine learning models like SVM and neural networks to create flexible product segmentation systems that improved inventory and service performance. Their study suggests that combining multiple data sources and techniques in layered, hybrid models can greatly improve segmentation effectiveness and business impact.

Overall, the research shows that modern e-commerce customer segmentation benefits from combining machine learning's power with decision-support systems and explainable analytics.

This mix of disciplines helps create customer groups that are useful for targeted marketing, product recommendations, and improving customer engagement. However, there are still challenges, like making models more transparent, handling large amounts of data, and ensuring they are scalable. Future research is looking more into explainable AI, real-time models, and combining different types of data to further improve segmentation and deliver personalized experiences that help businesses grow.

III. METHODOLOGY OF THE SYSTEM

The approach used for this e-commerce customer segmentation project is organized into several steps to identify clear customer groups that can help with targeted marketing.

The process includes data gathering, preparation, feature selection, normalization, clustering, visualization, and exporting insights. Each step is carefully designed to make sure the analysis is accurate, the data stays consistent, and the results can be used effectively in marketing strategies.

The first step is gathering the data.

A	B	C	D	E	F
ustomer_	ustomer_Nam	end_Amou	rchased_Discou	Timestamp	Segment
1	Sohel shaikh	130	TRUE	2025-09-24 19:11:21	Value Seekers
2	john	240	FALSE	2025-09-24 19:18:43	Premium Customers
3	brooklyn	150	FALSE	2025-09-24 19:18:59	Budget Shoppers
4	Lina	60	TRUE	2025-09-24 19:51:23	Value Seekers
5	Pradeep	170	FALSE	2025-09-25 11:33:33	Budget Shoppers
6	Sujal	80	TRUE	2025-09-25 12:33:12	Value Seekers
7	Dean	100	TRUE	2025-09-25 12:33:33	Value Seekers
8	Anderson	180	TRUE	2025-09-25 12:33:59	Value Seekers
9	KRISHNA	10	TRUE	2025-09-25 12:34:17	Value Seekers
10	Ghost	130	FALSE	2025-09-25 12:34:32	Budget Shoppers
11	HArsh	150	TRUE	2025-09-25 13:25:16	Value Seekers
12	John Smith	250	FALSE	2025-10-12 19:24:14	Premium Customers
13	Maria Garcia	89.5	TRUE	2025-10-12 19:24:14	Value Seekers
14	David Johnso	150	TRUE	2025-10-12 19:24:14	Value Seekers
15	Sarah William	45	FALSE	2025-10-12 19:24:14	Budget Shoppers
16	Michael Brow	300	FALSE	2025-10-12 19:24:14	Premium Customers
17	Emily Davis	75	TRUE	2025-10-12 19:24:14	Value Seekers
18	Robert Miller	180	FALSE	2025-10-12 19:24:14	Budget Shoppers
19	Lisa Andersor	95	TRUE	2025-10-12 19:24:14	Value Seekers

Fig-1 : Dataset

We collect information about customers, including their names, how much they spend, and whether they use discounts. This information is stored in a file called "basic_customers.csv" which starts the analysis. It's important that the data is accurate, complete, and representative, as this affects how well the segments are formed. In more complex systems, data might also come from other sources like past purchases, customer feedback, and website activity. Next, we prepare the data to make sure it's clean and ready for analysis.

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 5, Issue 2, November 2025
Impact Factor: 7.67

This involves checking for errors, filling in missing values, and converting some data into a format that can be used in calculations. For example, if a customer used a discount, it's marked as True or False, and we change that to a number. We also remove any duplicate entries and look for extreme values that might skew the results. This helps ensure that the data we use for the clustering is correct and consistent.

Then, we select and create features that best show customer behaviour. In this project, we focused on two main features: total spending and whether the customer used discounts. These two factors help us understand how customers spend money and how sensitive they are to prices. Depending on the project, other factors like how often customers shop, how recently they bought, or the average value of their orders can also be considered to enrich the segmentation. After selecting the features, we normalize the data to make sure each feature contributes equally to the clustering process.

```python
class CustomerSegmentation:

    def load_data(self):
        """Load existing customer data or create new file"""
        try:
            self.df = pd.read_excel(self.data_file)
            if 'Timestamp' in self.df.columns:
                self.df['Timestamp'] = pd.to_datetime(self.df['Timestamp'])
        except FileNotFoundError:
            # Create initial dataframe structure
            self.df = pd.DataFrame(columns=[
                'Customer_ID', 'Customer_Name', 'Spend_Amount',
                'Purchased_Discount', 'Timestamp', 'Segment'
            ])
            self.save_data()

    def save_data(self):
        """Save data to Excel file"""
        self.df.to_excel(self.data_file, index=False)

    def add_customer(self, name, spend, discount):
        """Add new customer to the dataset"""
        new_id = len(self.df) + 1 if len(self.df) > 0 else 1
        timestamp = datetime.now()

        new_customer = {
            'Customer_ID': new_id,
            'Customer_Name': name,
            'Spend_Amount': float(spend),
            'Purchased_Discount': bool(discount),
            'Timestamp': timestamp,
            'Segment': 'New'  # Temporary segment, will be updated after clustering
```

Fig-2: Code

This is important because spending amounts can be much larger than binary features like discount usage. By scaling all features to have the same range, we prevent any one feature from having too much influence. This allows the K-means algorithm to identify meaningful groups without bias. The clustering phase is where the actual grouping happens.

We use the K-means algorithm to divide customers into different groups based on their spending and discount usage. The number of groups, or clusters, is determined using the Elbow Method, which finds the point where adding more clusters doesn't significantly improve the results. In this project, the number of clusters was between two and four, depending on how the data looked. K-means works by randomly placing cluster centers in the data space.

Each customer is then assigned to the nearest center based on distance. The centers are updated by finding the average of all the customers in each cluster, and this process repeats until the clusters stop changing. The final groups represent customer segments with similar spending and discount behavior. Once the clusters are formed, we interpret them to give them meaningful names.

For example, customers who spend a lot and don't use discounts are labeled as "Premium Customers," while those who spend less and often use discounts are called "Value Seekers" or "Budget Shoppers." These labels help marketing teams understand the customer groups and plan targeted strategies. Finally, we visualize the results using tools like matplotlib and Plotly.

We create scatter plots to show how the clusters look in terms of spending and discount use, pie charts to show the proportion of each group, and interactive dashboards to explore the data in more detail. The clustered data, along with information on average spending, discount usage, and other metrics, is then exported as a CSV file for use in marketing systems or customer relationship management tools.

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 5, Issue 2, November 2025 Impact Factor: 7.67

A. Workflow

Data Collection: Collect customer transaction data, including how much they spent and which discounts they used.
Data Preprocessing: Fix any missing data and adjust numbers so they are easier to work with.
Feature Selection: Choose spend amount and discount usage as the main factors for grouping customers.
Clustering: Use K-means clustering, and decide the number of groups based on the data.
Segment Labelling: Look at the key traits of each group to give them clear names.
Visualization: Create charts to show the customer groups in an easy-to-understand way.
Export Results: Keep the findings ready for use in marketing strategies.

B. Algorithm (Pseudocode)

Customer data includes Spend Amount and Purchased Discount.
Output is Customer Segments.
Load the customer data.
If there are fewer than 3 customers, assign a default segment and stop.
Make sure the Spend Amount and Purchased Discount values are standardized.
Choose K based on how big the dataset is, usually between 2 and 4.
Use K-means clustering with K groups.
For each group center:
Look at the spend score and discount score.
Give a segment name based on set limits.
Give each customer a group label.
Save the data with the group labels.

C. Flowchart

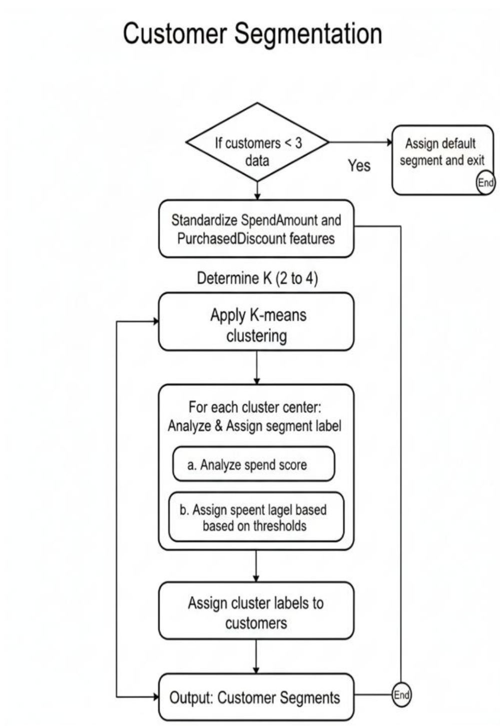


Fig-3 : Flowchart

IV. IMPLEMENTATION

The e-commerce customer segmentation project was built using a Streamlit application that makes it easy for marketers and data scientists to interact with the data, manage it, and get real-time insights.

The front end of the app uses Streamlit's simple tools, giving users an easy way to input information. They can add new customers one by one by filling out a form with details like names, spending amounts, and discount preferences. For handling multiple customers at once, the app lets users upload CSV files, making it quick to import and add more customer data. The interface also shows helpful messages and loading indicators so users know when data is being processed, making the experience smoother.

On the back end, the app is made with Python and uses powerful data science tools like pandas for managing data and scikit-learn for running K-means clustering.

This part of the app cleans and prepares the data—like fixing errors, changing data types, and picking useful features—so it works well with the clustering algorithm. Every time new data comes in, the clustering logic runs automatically, and customer groups are updated in real time. The back end also connects well with the front end, making sure the whole system stays responsive.

For showing results, the app uses Plotly to create interactive scatter plots that show how customers are grouped based on spending and discount use, with different colours for each group.

This helps users quickly see how the segmentation is working. Other charts like pie charts show the share of customers in each group, while tables list detailed customer information along with their assigned segment and behavior. These tools help users explore data in a clear and useful way. To help users learn and test, the app includes sample datasets they can download directly.

These examples show the right data format and let users try out segmentation steps easily, making it easier to understand and start using the tool. The app also has real-time dashboards that show important metrics like the total number of customers, average spending in each group, how much discounts are used, and the number of customer groups.

These numbers update as data changes, so users can track trends and changes in their customer base instantly.

Finally, the app lets users export all their segmentation results, including customer details, group labels, and summary statistics, in CSV or Excel formats. This makes it easy to use the data with other tools like CRM systems, marketing software, and offline analysis programs, helping with decisions that go beyond what the app can do.

V. RESULTS AND ANALYSIS

The results from using K-means clustering on the e-commerce customer dataset show clear and meaningful groups that match common marketing categories. The sample had 10 customer records with different spending amounts and discount usage, which allowed the clustering algorithm to identify distinct customer behavior patterns. The clustering process found several easy-to-understand groups.

One group includes high spenders who rarely use discounts, called "Premium Customers." These customers have strong purchasing power and care more about quality or convenience than deals. Another group is "Value Seekers," who often use discounts, showing they are sensitive to promotions. The third group, "Budget Shoppers," consists of customers with lower spending and little discount use, usually price-conscious buyers. These separate groups give useful insights for creating focused marketing messages and custom campaigns that match each group's preferences.

Visual tools like scatter plots that show spending versus discount usage and pie charts that summarize segment sizes help confirm the accuracy of the segmentation.

The scatter plot clearly shows the different groups based on their behavior, while the pie chart gives a quick view of how large each segment is, which helps in deciding how to allocate resources. Together, these tools make the results easier to understand and help engage stakeholders.

An important aspect is that the system can handle new data added over time, allowing clusters to change and reflect new customer behaviours. This helps marketers spot new groups or changes in customer trends quickly, making marketing efforts more flexible and responsive.

When compared to existing research, the findings align with many studies that support K-means as an effective method for customer segmentation in e-commerce.

K-means is known for its simplicity and ability to clearly separate customers based on their behaviours, which helps in creating better-targeted marketing, personalized recommendations, and improved return on investment. This agreement shows the method is both practical and backed by academic research, even when working with limited data, and provides useful business insights.

Overall, the findings show that using K-means clustering for targeted customer segmentation can improve marketing by tailoring offers and communications to specific customer needs. This approach helps improve customer acquisition, engagement, and long-term value, supporting the shift from broad marketing tactics to more precise strategies that are essential for staying competitive in the changing e-commerce environment.

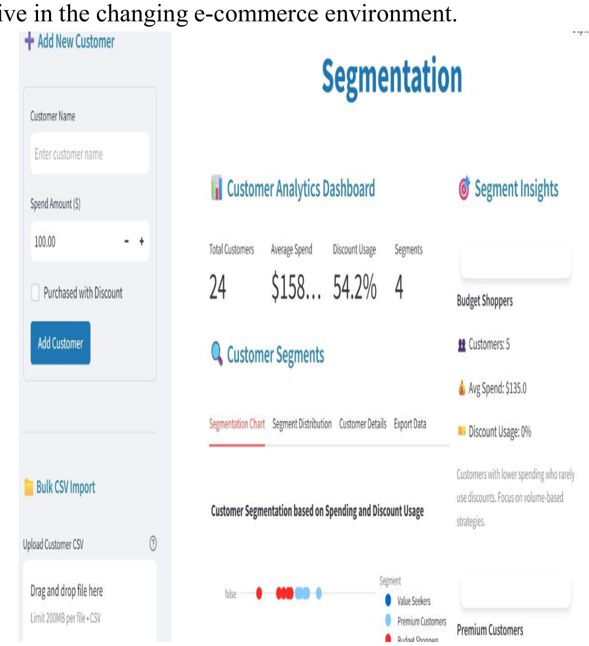


Fig-4 : Interface

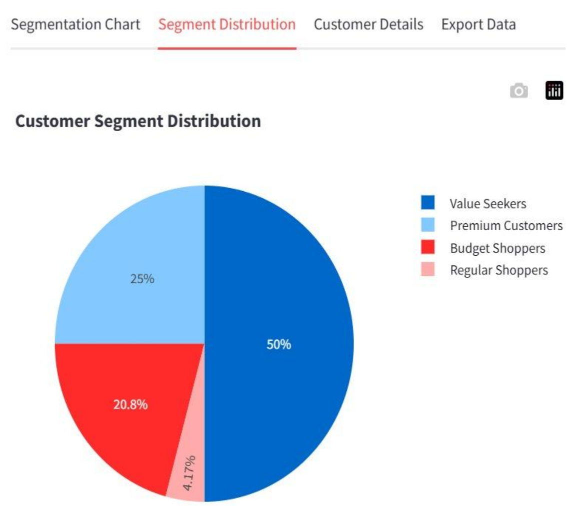
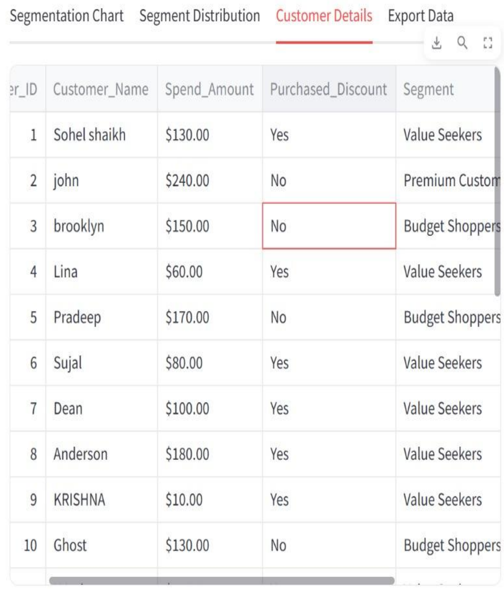


Fig-5 : Customer Distribution chart

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 5, Issue 2, November 2025
Impact Factor: 7.67


Segmentation Chart Segment Distribution Customer Details Export Data

er_ID	Customer_Name	Spend_Amount	Purchased_Discount	Segment
1	Sohel shaikh	$130.00	Yes	Value Seekers
2	john	$240.00	No	Premium Custom
3	brooklyn	$150.00	No	Budget Shoppers
4	Lina	$60.00	Yes	Value Seekers
5	Pradeep	$170.00	No	Budget Shoppers
6	Sujal	$80.00	Yes	Value Seekers
7	Dean	$100.00	Yes	Value Seekers
8	Anderson	$180.00	Yes	Value Seekers
9	KRISHNA	$10.00	Yes	Value Seekers
10	Ghost	$130.00	No	Budget Shoppers

Fig-6: Customer Details Csv

VI. FUTURE SCOPE

The future of using data science and K-means clustering for e-commerce customer segmentation offers many exciting possibilities that can greatly improve marketing accuracy, how well customers are engaged, and overall business results. One big chance is to include more customer details than just spending and using discounts.

Adding factors like how often someone shops and how recently they bought can help group customers based on their loyalty and where they are in their buying journey. Looking at how customers behave online—like how they browse websites, how long they stay on pages, how often they leave items in their cart, and how they use mobile apps—gives a more complete picture of their habits. These extra details help deeply understand what customers want and prefer, leading to better personalized offers. For example, knowing if someone is just looking around or actually buying can help create different kinds of promotions. Using more advanced clustering methods can also improve how well customers are grouped.

While K-means is good for its simplicity, other techniques like hierarchical clustering, DBSCAN, or models based on neural networks (like self-organizing maps or autoencoders) can find more complex patterns, spot unusual data, and detect clusters with irregular shapes. These tools can uncover subtle customer groups, niche markets, and new trends that simpler methods might miss. Another promising area is making models that can adapt in real-time.

Instead of only updating customer groups at certain times, real-time models keep changing as new data comes in. This helps track changing customer behavior, seasonal changes, and the effects of marketing campaigns, allowing marketers to quickly adjust offers and messages. Tools like Streamlit or other dashboard platforms can show live results from these changes. Expanding segmentation to cover all customer channels, not just one, helps create more personalized experiences.

By bringing together data from physical stores, websites, social media, apps, and customer service calls, businesses can build full profiles of their customers. This broader view helps make segments more accurate and supports a consistent experience across all touchpoints. For example, knowing someone spends a lot on their phone but uses coupons in stores can lead to tailored offers that work across all channels. Combining segmentation with recommendation engines and marketing automation also works well.

Segments can help AI systems suggest products that match what each group likes and has bought before. Automated campaigns can then send personalized emails, app notifications, or ads based on these groups, making marketing more effective at scale. This full integration helps improve the customer journey and conversion rates.

Finally, using predictive analytics and machine learning systems can help businesses forecast customer needs, the risk of customers leaving, their long-term value, and how likely they are to respond to offers.

When combined with customer segments, these predictions let businesses act before problems arise and seize opportunities for upselling or keeping customers with the best use of their resources.

VII. CONCLUSIONS

This study clearly shows how the K-means clustering algorithm works well for grouping e-commerce customers based on two main factors: how much they spend and how often they use discounts. This method helps create clear customer groups like high-spending customers, budget-conscious buyers, and those who look for deals, each with their own shopping habits and how they respond to marketing.

The approach allows e-commerce companies to get useful information that helps them tailor marketing efforts more effectively. By understanding what each group likes, businesses can create better promotions and use marketing resources more wisely, which can lead to higher customer involvement, more sales, and more income. The framework can adapt to different types of e-commerce data, making it useful for datasets that vary in size and makeup.

In addition, the system has a simple and easy-to-use interface through an interactive Streamlit dashboard, making it practical for marketing teams who may not have advanced data science skills. This ease of use and ability to scale help more businesses adopt the tool, allowing them to keep improving their marketing strategies as customer data changes over time.

In general, K-means clustering is a popular method for customer grouping because it's easy to understand, works quickly, and has been successfully used in real-world marketing situations, as seen in retail and business-to-business areas. Using spending and discount behavior as the key factors for clustering captures important aspects like how much customers value money and how sensitive they are to prices, which are important for making smart marketing choices.

To sum up, this research shows that using K-means clustering for customer grouping in e-commerce delivers real business benefits through better personalization and extracting more customer value. Its flexible design and simple use make it a useful tool for improving marketing performance in ever-changing e-commerce settings, helping with better customer relationships and supporting long-term business success.

VIII. ACKNOWLEDGEMENT

The authors would like to thank the open-source software communities for their important contributions. The tools and libraries from these communities were essential to this research. Python's wide range of tools, especially the panda’s library for handling data, scikit-learn for machine learning, and visualization tools like Plotly and matplotlib, played a key role in making the development process efficient, supporting experimentation, and producing meaningful results.

We also want to express our appreciation to the global data science and e-commerce analytics researchers. Their previous work, publications, and datasets greatly influenced the methods and decisions made during this project. Their research on using machine learning for customer segmentation helped shape the clustering techniques and visualization approaches used here.

Our thanks also go to the reviewers who provided helpful feedback and thorough evaluations. Their input significantly improved the quality, clarity, and depth of the research. Their expertise helped refine the overall structure of the study and made the analysis more solid.

Lastly, we acknowledge the dataset providers who made real customer data available for testing and validation. Access to real-world data was crucial in showing how K-means clustering can be effectively used for customer segmentation and targeting in e-commerce.

These efforts from open-source developers, researchers, reviewers, and data contributors all played a vital role in making this research possible.

Their contributions have helped move forward data-driven marketing strategies in the e-commerce industry.

REFERENCES

[1]. Alamsyah, A., & Perbangsa, A. S. (2021). Customer Segmentation Based on Transaction Data Using K-Means and RFM Model. *Indonesian Journal of Information Systems.* https://ijis.org/index.php/ijis/article/view/255

[2]. Siami-Namini, S., & Namin, A. S. (2019). A Comparative Study of Time Series Forecasting of Sales Using LSTM.
arXiv. https://arxiv.org/abs/1909.10549

[3]. Wang, H., & Wang, N. (2020). Research on E-commerce Customer Segmentation Based on K-means Clustering Algorithm. *Journal of Physics: Conference Series.* https://iopscience.iop.org/article/10.1088/1742-6596/1574/1/012028

[4]. Kumar, V., & Reinartz, W. (2018). Customer Relationship Management: Concept, Strategy, and Tools. *Springer.* https://link.springer.com/book/10.1007/978-3-662-55381-7

[5]. Sarstedt, M., & Mooi, E. (2019). A Concise Guide to Market Research: The Process, Data, and Methods Using R.
Springer. https://link.springer.com/book/10.1007/978-3-662-56707-4

[6]. Huang, M., Wang, H., & Wang, Z. (2021). A Hybrid Recommendation System Based on Customer Segmentation.
IEEE Access. https://ieeexplore.ieee.org/document/9397854

[7]. Li, Y., & Kauffman, R. J. (2022). Customer Segmentation Using Unsupervised Learning: A Data-Driven Framework.
Journal of Management Analytics. https://www.tandfonline.com/doi/full/10.1080/23270012.2022.2046542

[8]. Li, F., & Xu, X. (2020). E-commerce Customer Segmentation Using Machine Learning. *International Journal of Advanced Computer Science and Applications.* https://thesai.org/Publications/ViewPaper?Volume=11&Issue=7&Code=IJACSA&SerialNo=54

[9]. Ruchi Rani, et al. "Customer segmentation in e-commerce: K-means vs hierarchical clustering," TELKOMNIKA, 2025. URL: https://telkomnika.uad.ac.id/index.php/TELKOMNIKA/article/view/26384

[10]. PLOS ONE. "Customer segmentation in the digital marketing using a Q-learning-based approach," 2025. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0318519

[11]. Bhatia, Sinha et al. "Targeted customer classification using machine learning," WJARR, 2025. URL: https://journalwjarr.com/sites/default/files/fulltext_pdf/WJARR-2025-0334.pdf

[12]. "Customer Segmentation - IJRASET," 2025. URL:https://www.ijraset.com/best-journal/customer-segmentation

[13]. Scientific Reports, "A dynamic customer segmentation approach by combining LRFMS and time series clustering," 2024.URL:https://www.nature.com/articles/s41598-024-68621-2

[14]. "A Review of Customer Segmentation Methods," IJARSCT, 2025. URL: https://ijarsct.co.in/Paper8904.pdf