Customer Segmentation

Team Members: Gayathri Mummaneni, Uma Viraja Dasaraju
University Of Missouri – Kansas City

Abstract

The goal of modern organizations is to provide their clients with highly customized services, which is only feasible with some sort of consumer segmentation or classification. By targeting their clients, businesses can easily arrange their services and products around them and increase income Customer segmentation is a crucial technique in data-driven business intelligence, enabling companies to understand customer behavior and improve marketing strategies. This study explores various clustering techniques, including K-Means, Hierarchical Clustering, Gaussian Mixture Model (GMM), and Fuzzy C-Means (FCM), to categorize customers based on their spending habits and demographics. By leveraging mall customer data, we identified five distinct customer segments, each representing unique spending behaviors. Our findings highlight the effectiveness of machine learning-based segmentation in enhancing targeted marketing, improving customer retention, and optimizing business operations. The study also compares clustering performance using silhouette scores and discusses future improvements, such as real-time segmentation and deep learning-based approaches. These insights provide a strong foundation for businesses to develop data-driven marketing strategies that cater to specific customer needs.

Introduction

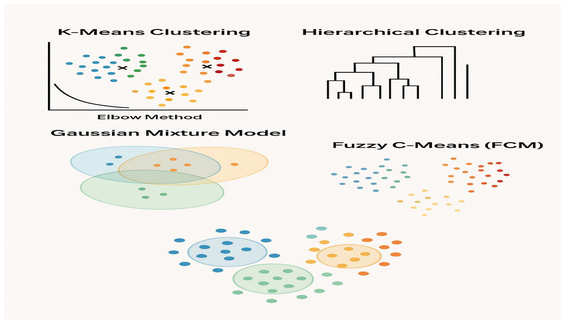
Customer segmentation is vital for business intelligence, allowing companies to target products and marketing more effectively. Using mall customer data, including spending behavior and demographics, we implemented clustering techniques to identify distinct customer segments. By segmenting customers, businesses can enhance customer retention, increase profitability, and improve customer satisfaction. Traditional segmentation methods rely on heuristics, but data-driven clustering techniques provide deeper insights, revealing hidden patterns in consumer behavior. This study focuses on leveraging clustering algorithms to develop meaningful segments that aid in personalized marketing strategies.

Objective

•Identify distinct customer groups based on spending patterns and demographics to enhance business intelligence.
•Compare and evaluate various clustering algorithms to determine optimal segmentation strategies.
•Provide actionable insights that businesses can leverage to implement targeted marketing campaigns and customer engagement strategies.
•Establish a framework for automated and scalable customer segmentation using machine learning techniques.
•Improve customer experience by tailoring services and promotions to specific demographic and behavioral segments.

Methodology

We explored four clustering techniques to segment customers based on their shopping behaviors:
•**K-Means Clustering**: Applied the Elbow method to determine the optimal number of clusters. This method minimizes variance within clusters, ensuring meaningful segmentation.
•**Hierarchical Clustering**: Used agglomerative clustering to build a hierarchy of clusters, visualized using a dendrogram. This technique is advantageous for understanding relationships between clusters but can be computationally expensive.
•**Gaussian Mixture Model (GMM)**: A probabilistic clustering technique that allows data points to belong to multiple clusters. This method provides flexibility in handling overlapping clusters, making it suitable for complex data distributions.
•**Fuzzy C-Means (FCM)**: Unlike traditional clustering, FCM assigns probabilities to data points belonging to multiple clusters. This soft clustering approach is useful when customer behaviors are not strictly categorical.
Each method was assessed using the **silhouette score**, which evaluates cluster cohesion and separation to determine the effectiveness of segmentation.



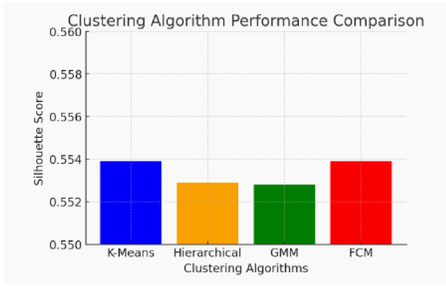
Future Results

The analysis led to the identification of five customer clusters with distinct spending behaviors:
•**Cluster 0**: Moderate income ($39K-$80K), moderate spending scores. Represents customers who balance spending and saving habits.
•**Cluster 1**: High income ($68K-$137K), low spending scores. These customers have substantial financial capacity but exhibit conservative spending behavior.
•**Cluster 2**: Low income ($10K-$40K), low spending scores. This segment consists of budget-conscious shoppers who prioritize savings.
•**Cluster 3**: Low-moderate income ($10K-$39K), high spending scores. These customers demonstrate aspirational buying behavior, prioritizing premium purchases despite limited incomes.
•**Cluster 4**: High income ($68K-$137K), high spending scores. These affluent shoppers spend significantly, making them prime targets for premium services.

Performance Evaluation

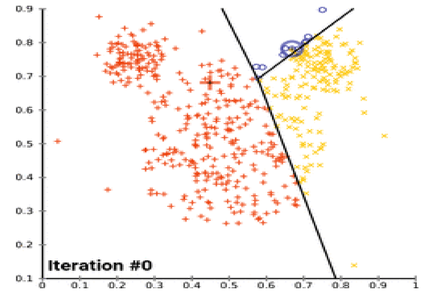
Algorithm	Silhouette Score
K-Means	0.5539
Hierarchical	0.5529
GMM	0.5528
FCM	0.5539

The results indicate that K-Means and FCM performed slightly better, offering clear and distinct cluster separation. The choice of method depends on the business use case, with GMM being preferable for complex overlapping behaviors and hierarchical clustering providing better visualization of customer relationships.



Conclusion

Segmenting customers based on income and spending helps businesses tailor marketing strategies for better engagement. For example:
•**Cluster 0 (Moderate Spenders)**: Value-focused promotions, discounts, and loyalty programs.
•**Cluster 1 (Affluent Low Spenders)**: Upselling opportunities through personalized incentives and exclusive offers.
•**Cluster 2 (Budget-Conscious Shoppers)**: Promotions focusing on affordability and essential products.
•**Cluster 3 (Aspirational Shoppers)**: Luxury product marketing and aspirational branding strategies.
•**Cluster 4 (Affluent High Spenders)**: Exclusive membership programs, premium service offerings, and personalized experiences.
By leveraging these data-driven insights, businesses can improve customer targeting, boost conversion rates, and maximize revenue.



Recommendations

1.Lalwani, P., et al. "Customer churn prediction system: A machine learning approach," *Computing*, 2021.
2. Nandapala, E. Y. L., et al. "The practical approach in customer segmentation using K-Means," *IEEE ICIIS*, 2020.
3.Tressa, N., et al. "Customer-based market segmentation using clustering in data mining," 2024.
4.Toh, W. X. "Customer segmentation on clustering algorithms," *eprints.utar.edu.my*, 2023.

Acknowledgements

We extend our gratitude to our advisors and mentors at the University of Missouri – Kansas City for their guidance and support. Additionally, we acknowledge the contributions of various researchers whose work provided the foundation for our study.