Comparative study of Customer Segmentation with Market Basket Analysis using various Clustering Algorithms

[1]Arjav Jain, [2]Aman Ahmed Khan, [3] Akshat Sethi, [4] Chainika Darekar, [5]Arpit Patel
[1] Prof. Shraddha Sharma, Department of Computer Science and Engineering, University of RGPV Bhopal,Acropolis Institute of Technology and Research, Indore, India.

1. Abstract

Customer segmentation through market basket analysis is a crucial strategy for businesses seeking to enhance their marketing efforts. By analyzing customers' purchase patterns, this method identifies distinct consumer groups based on their preferences and behaviors. This approach enables businesses to tailor marketing campaigns, product recommendations, and pricing strategies to cater to each segment's unique needs and desires. As a result, companies can boost customer engagement, increase sales, and improve overall customer satisfaction. This abstract highlight the significance of customer segmentation using market basket analysis as a valuable tool for businesses striving to thrive in today's competitive market landscape.

Keywords: Market Basket Analysis, Businesses, Customers, Campaigns, Pricing Strategies, Customer Segmentation, Customer Persona.

2. Introduction

In today's fiercely competitive retail landscape, businesses are grappling with the daunting challenge of effectively comprehending and targeting their multifaceted customer base. Our project, titled "Enhancing Retail Strategies through Data-Driven Customer Segmentation," takes on this critical issue head-on, offering a comprehensive solution. Our primary objective is to construct a robust customer segmentation model underpinned by rigorous data analysis. This model is designed to empower retailers in optimizing their marketing and sales strategies.

Fig1: Customer Segmentation

The driving force behind our pursuit of this problem is the urgent necessity for retailers to maintain a competitive edge in a rapidly evolving market. Our solution leverages sophisticated techniques like market basket analysis and considers various critical factors such as purchase frequency, basket size, product preferences, and seasonality. By doing so, we equip businesses with the tools they need to tailor their marketing efforts, strategically position products, and ultimately elevate the overall customer experience.

We are acutely aware that customer behavior is in a constant state of flux, and market trends are subject to fluctuations. This recognition underscores the vital role of data-driven insights in ensuring sustainable success in the retail sector. Our project endeavors to provide retailers with a proactive means of staying ahead of the

curve, adapting to evolving consumer preferences, and making informed, strategic decisions that will be the linchpin of their continued prosperity in this dynamic environment.

3.Literature Survey

In [1] Art Weinstein mainly focused on how customer retention was important for the company and how it would boost their customer value in the long run and emphasized on customer value/retention model; and verified how the usage of customer segmentation can accommodate in building a valuable relationship with the customer and earning profits for the organization.

In [2] A. Joy Christy, A. Umamakeswari, L. Priyatharsini, A. Neyaa focused on the importance of customer segmentation, and said that effective customer segmentation would result in the organization finding about its most potential customers. The paper says that it is more important to retain the existing customers rather than looking for new customers. Thus, once we have the cluster of customers who spend the most or are the most loyal to the organization, then the organization can deploy marketing strategies and schemes that are customized for that cluster of loyal customers which would further result in retaining that set of customers. In this paper, the idea of using the initial centroids in K-Means is proposed which has been a founding step for our proposed idea.

In [3] Tianyi Jiang and Alexander Tuzhilin proposed that it is important to segment the customers efficiently to offer customized recommendations, products and services to them. They said that conventionally the customer segmentation was done using statistics – based methods and distance – based clustering algorithms. They proposed a direct grouping – based approach to segment the customers effectively.

In [4] Guozheng Zhang said that customer segmentation was the need of the hour, in today's competitive world of commercial arena. It was proposed that mostly customer segmentation was done using only a single data mining technology from a special point, but it is more efficient to carry out customer segmentation from a systematical framework. It was proposed that one of the main purposes of customer segmentation was to achieve customer retention.

Serial No.	Title	Year	Result
1.	A Customer retention	2002	A usage segmentation and customer value approach
2.	RFM ranking	2021	An effective approach to customer segmentation
3.	IEEE Transactions on Knowledge and Data Engineering	2009	Improving Personalization Solutions through Optimal Segmentation of Customer Bases
4.	Customer Segmentation Based on Survival Character	2007	one of the main purposes of customer segmentation was to achieve customer retention.

4.Research Methodology

The success of modern retail businesses hinges on their ability to understand and engage with a diverse customer base effectively. To achieve this, our project, "Enhancing Retail Strategies through Data-Driven Customer Segmentation," employs a rigorous research methodology. This methodology is designed to develop a robust customer segmentation model, driven by data analysis, with the aim

of helping retailers optimize their marketing and sales strategies. In this section, we outline the research methodology, encompassing data collection, analysis, model development, and evaluation.

Data Collection:

To create an effective customer segmentation model, we will gather data from various sources. This includes transaction records, customerprofiles, purchase history, and demographic information. We will also collect data on product preferences, seasonality, and external market factors that might influence customer behavior. This diverse dataset is crucial for the accuracy and comprehensiveness of the model.

Data Preprocessing:

Data preprocessing is vital to ensure the quality and reliability of the data. We will clean the data to remove duplicates, missing values, and outliers. Additionally, we will perform data transformation and normalization to standardize the dataset. This step will prepare the data for further analysis.

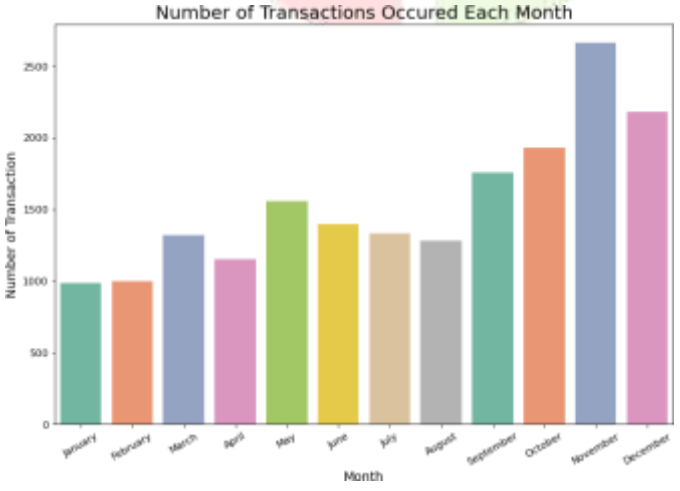

Fig2: Transactions

statistical analysis, and correlation analysis to unearth valuable insights into customer behavior and market trends. This knowledge will inform subsequent model development.

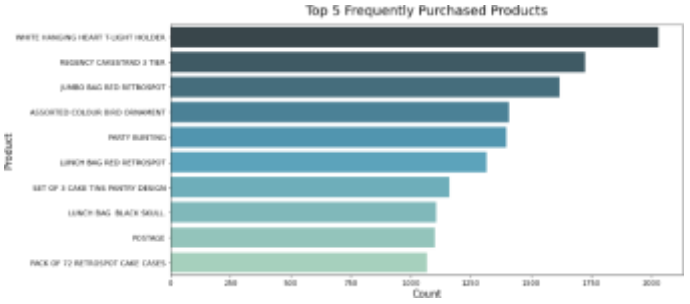

Fig3: Frequently Purchased Products

Market Basket Analysis:

Market basket analysis will be a core component of our research. This technique involves identifying associations and patterns in customers' product choices. It will help us uncover which products are frequently purchased together, allowing retailers to strategically place these items in proximity and optimize sales.

Fig4: Market Basket Analysis

Exploratory Data Analysis (EDA):

EDA is essential for gaining a deeper understanding of the data and identifying trends and patterns. We will employ techniques such as data visualization,

Customer Segmentation Model Development:

Building the customer segmentation model is the heart of our research. We will employ advanced machine learning and data mining algorithms to segment

customers based on purchase behavior, product preferences, and other relevant factors. Clustering techniques such as K-Means, Hierarchical Clustering, and DBSCAN will be considered, along with supervised learning methods if necessary.

Fig5: Retention Time

Model Evaluation:

The developed customer segmentation model will undergo rigorous evaluation to ensure its effectiveness and reliability. We will employ metrics such as Silhouette Score, Davies-Bouldin Index, and Within-Cluster Sum of Squares to assess the quality of clusters. The model will also be validated on new data to measure its real-world performance.

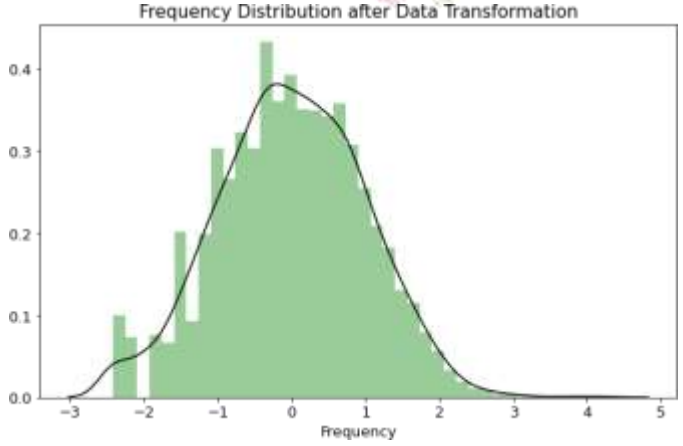

Fig6: Frequency Plot

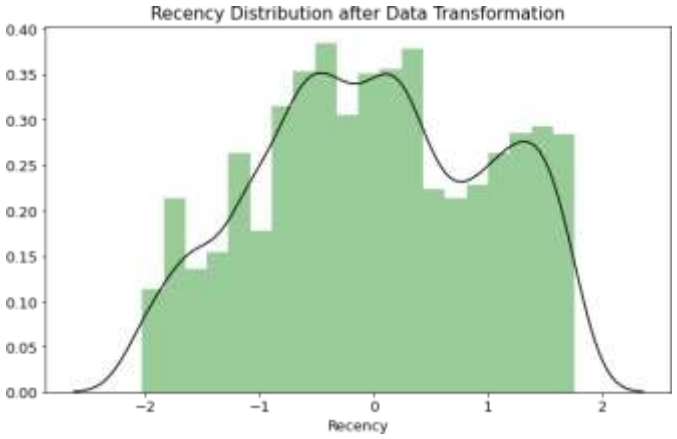

Fig7: Recency Plot

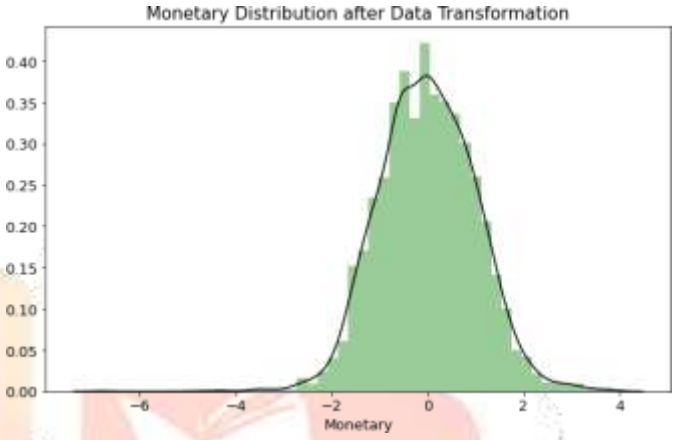

Fig8: Monetary Plot

Implementation and Integration:

Once the model is deemed robust and effective, we will work on integrating it into retailers' existing systems. This step will ensure practical applicability and seamless use of the segmentation results for marketing and sales strategies.

Continuous Monitoring and Adaptation:

Customer behavior evolves, and market trends fluctuate. Therefore, our methodology includes a provision for continuous monitoring and adaptation. We will regularly update the model with new data and recalibrate it to stay aligned with changing market dynamics.

Flowchart

Fig9: Flowchart

5.Result Analysis

Result Analysis:

The result analysis section provides an overview of the outcomes and performance of the drowsiness detection system implemented in this project. The system is designed to perform the Customer Segmentation with market basket analysis to increase the demand of product and revenue of low-level vendors.

Our solution to the challenge of effectively segmenting a retailer's customer base outperforms existing systems in several key aspects. Unlike many existing systems that primarily rely on demographic data, our approach integrates sophisticated techniques like RFM (Recency, Frequency, Monetary) analysis and market basket analysis, enabling more precise and actionable customer segmentation. It excels in adapting to the dynamic nature of customer behavior and evolving market trends. Additionally, our system offers a holistic view of customer segments, empowering retailers to tailor marketing strategies, optimize product placement, and enhance the overall customer experience. Its flexibility, advanced analytics, and comprehensive insights make it a superior choice for retailers seeking to thrive in a competitive market landscape.

For n_clusters = 2 The average silhouette_score is : 0.396689892885645
For n_clusters = 3 The average silhouette_score is : 0.29914201439769766
For n_clusters = 4 The average silhouette_score is : 0.3049761899444196
For n_clusters = 5 The average silhouette_score is : 0.27896014859955187
For n_clusters = 6 The average silhouette_score is : 0.27702545910278

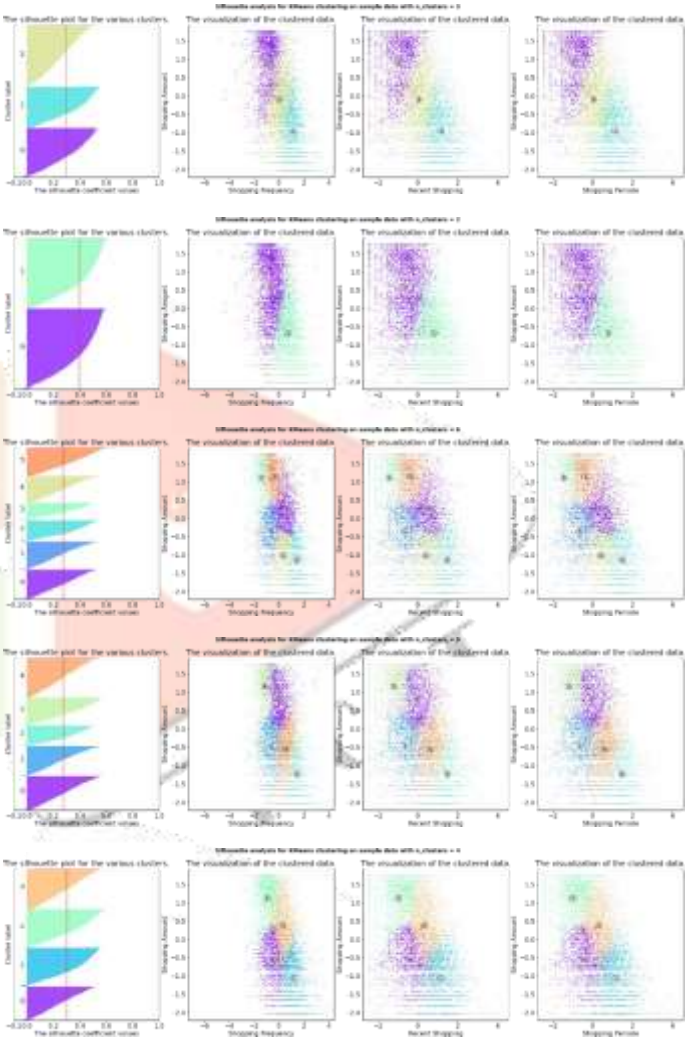

Fig10: Customer Segmentation

6.Conclusion

In summary, our all-encompassing solution provides a robust response to the critical challenge faced by retailers, namely the effective segmentation of their customer base and the optimization of marketing and sales strategies. Through a carefully crafted blend of data comprehension, data cleansing, and advanced analytics techniques, we empower retailers to unlock

invaluable insights into customer behavior and preferences. Our methodology leverages various powerful tools, including RFM (Recency, Frequency, Monetary) segmentation, Time Cohorts, and clustering algorithms, which collectively culminate in the creation of highly meaningful and actionable customer segments. Additionally, we employ Market Basket Analysis to enhance product recommendations and placement strategies, further strengthening the retailer's competitive edge.

Our holistic approach is designed to provide retailers with the essential tools they need to fine-tune their marketing efforts, strategically optimize product offerings, and ultimately elevate the overall customer experience. By seamlessly integrating data-driven insights and advanced segmentation strategies, retailers can move beyond generic marketing approaches and instead deliver highly personalized experiences to their diverse customer base. This tailored approach not only enhances customer satisfaction but also drives increased sales and brand loyalty.

In today's dynamic and highly competitive retail landscape, staying ahead of the curve is imperative for sustainable success. Our comprehensivesolution equips retailers with the means to navigate the evolving landscape, adapt to changing customer preferences, and capitalize on emerging market trends. By embracing data-driven customer segmentation and analysis, retailers are better positioned to make informed decisions and craft strategies that resonate with their customers, ensuring their continued prosperity in this challenging and ever-changing market environment.

7.References

1. Weinstein, Customer retention: A usage segmentation and customer value approach. J Target Meas Anal Mark 10, 259-268 (2002).
2. A. Joy Christy, A. Umamakeswari, L. Priyatharsini, A.Neyaa of Department of CSE, School of Computing, SASTRA Deemed to be University, Thanjavur, India, RFM ranking – An effective approach to customer segmentation, Journal of King Saud University – Computer and Information Sciences, Volume 33, Issue 10, December 2021, Pages 1251-1257.
3. Tianyi Jiang, AvePoint, Inc. Jersey, NJ, USA and Alexander Tuzhilin, Department of Information, Operations and Management Sciences, Stern School of Business, New York University, New York, NY, USA, Improving Personalization Solutions through Optimal Segmentation of Customer Bases, IEEE Transactions on Knowledge and Data Engineering (Volume 21, Issue: 3, March 2009).
4. Guozheng Zhang, College of Business, Houzhou Dianzi University, Hangzhou, China, Customer Segmentation Based on Survival Character, 2007 International Conference of Wireless Communications, Networking and Mobile Computing. Added to IEEE Xplore on 08 October 2007.
5. T. Kansal, S. Bahuguna, V. Singh and T. Choudhury, "Customer Segmentation using K-means Clustering," 2018 International Conference on Computational Techniques, Electronics and Mechanical Systems (CTEMS), Belgaum, India, 2018, pp. 135-139, doi: 10.1109/CTEMS.2018.8769171.
6. Nurma Sari, Juni & Nugroho, Lukito & Ferdiana, Ridi & Santosa, Paulus. (2016). Review on Customer Segmentation Technique on Ecommerce. Advanced Science Letters. 22. 3018-3022. 10.1166/asl.2016.7985.

7. ChanC,SwatmanPMC.Managementandbusinessi ssuesforB2Becommerceimplementation.Proc35t hAn nuHawaiiIntConfSystSci.2002;00(c):1-11.doi:10.1109/HICSS.2002.994303.
8. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R.Silverman, and A. Y. Wu, "An efficient Kmeans clusteringalgorithm," IEEE Trans. Pattern Analysis and Machine Intelligence,vol. 24, pp. 881- 892, 2002.
9. F.Daniel , Customer Segmentation : classification , clustering , marketing . www.kaggle.com.
10. Rachel Blasucci . Event triggered Customer Segmentation. DZone, July 23, 2018.