



IJIRCCCE

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 11, Issue 4, April 2023

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.379



9940 572 462



6381 907 438



ijircce@gmail.com



www.ijircce.com

Customer Mall Segmentation using K-means Algorithm

Sanmay Majalekar, Prem Hogade, Sourabh Kesharwani, Vishwajeet Powar, Dr. D.A.Nikam

B.E Student, Dept. of C.S., Dr. J. J. Magdum College of Engineering, Jaysingpur, India

B.E Student, Dept. of C.S., Dr. J. J. Magdum College of Engineering, Jaysingpur, India

B.E Student, Dept. of C.S., Dr. J. J. Magdum College of Engineering, Jaysingpur, India

B.E Student, Dept. of C.S., Dr. J. J. Magdum College of Engineering, Jaysingpur, India

HOD, Dept. of C.S., Dr. J. J. Magdum College of Engineering, Jaysingpur, India

ABSTRACT: In today's world, a vast amount of data is collected daily, and analyzing this data is crucial. For businesses to succeed in this era of innovation and intense competition, it's essential to have a business plan that aligns with current circumstances. To cater to the confusion of potential clients about what to buy and what not to buy, cutting-edge concepts are used in the work completed today. One such concept is customer segmentation, where clients are grouped based on similar characteristics. This process enables advertisers to focus their efforts on specific groups of clients, increasing the chances of a sale. It also allows them to use explicit communication channels to reach out to different clients and draw them in. The use of online media posts and postal letters to reach everyone is no longer practical. Adopting customer segmentation aids companies in developing stronger relationships with clients, which enhances the overall presentation of the company.

KEYWORDS: Target Customers, Clusters, Segmentation, Market Basket Analysis, K-means algorithm

I. INTRODUCTION

In the past, building and maintaining strong long-term client relationships relied on managing and maintaining customer interactions. However, in the modern era, the value of considering consumers as an organization's primary asset is increasing. Businesses are willing to invest in developing customer acquisition, maintenance, and development plans.

Customer segmentation is a process that helps to identify customers with different preferences, expectations, desires, and other characteristics. The primary objective of customer segmentation is to group individuals based on their shared interests, enabling the marketing team to develop an effective marketing strategy. Clustering is an iterative process that involves extracting knowledge from vast amounts of unstructured and raw data. Clustering is used in many fields, including exploratory data mining, machine learning, classification, and pattern recognition.

Customer segmentation refers to the practice of dividing a customer base into various groups or segments, based on shared traits. As mentioned in [2], these traits may include gender, age, interests, and various spending patterns that are relevant to marketing.

Segmentation is crucial as it enables businesses to customize their marketing strategies to suit each customer segment, make informed business decisions, identify products that appeal to each segment, and manage supply and demand for these products. Additionally, customer segmentation helps businesses to identify potential customers, predict customer defection, and find solutions to various customer-related challenges. Therefore, it is an essential practice for businesses that want to thrive in today's market.

II. LITERATURE SURVEY

According to. [1], "A Strategy for Targeted Customer Services" International Journal of Advanced Research in Artificial Intelligence (IJARAI), 4(10), 2015. – Papers purpose was to show how K-mean algorithm can be used for Segmentation.

According to [2]. New Strategy for Targeted Actions by Lahcen Abidar, Dounia Zaidouni, Abdeslam Ennouaary- Papers purpose was to propose a new model based on RFM model Recency, Frequency and monetary & k-mean algorithm to resolve those challenges.

According to [3], Paper focused on customer segmentation on their characteristics & attributes for E-commerce business.

According to [4] - In this paper ANOVA technique was carried out to test stability of clusters.

II. PROPOSED METHODOLOGY AND DISCUSSION

The process for customer mall segmentation that is being presented in this document consists of three stages.

Dataset: The mall will gather information about its customers' fundamental characteristics, such as Customer ID, age, gender, annual income, spending score, and email. The spending score can be assigned based on specific criteria such as customer behavior and purchase history.

Data Pre-processing (Data Cleaning): The input fields include Email, Customer ID, Gender, Age, and Annual Income. Customer ID is removed from the dataset for later processing. Gender data is converted to integer values using one-hot coding, such as Male=1 and Female=0. Records with empty emails are removed, and duplicates are deleted. Null values in the Age, Annual Income, and Spending Score columns are replaced with median values. Records with incorrect email addresses are removed, and only the Annual Income and Spending Score columns with values greater than 20 are selected.

Segmentation: Segmentation refers to dividing the mall's target market into groups of potential customers with similar demands and behaviors. This allows the business to market differently to each customer group according to their specific needs. The K-means clustering algorithm is used to calculate centroids, which iterate until the best centroid is found. It is assumed that the number of clusters is known, and the flat clustering algorithm is used. The 'K' in K-means stands for the number of clusters identified by the algorithm from the data.

Step 1: Specify the number of clusters, K, that the method should generate.

Step 2: Randomly select K data points and place each in a cluster.

Step 3: Compute the cluster centroids.

Step 4: Repeat previous steps until the ideal centroid is identified, which is determined by assigning data points to constant-size clusters.

Mailing: After data segmentation, the administrator will create offers, and corresponding emails will be sent to client

III. SIMULATION RESULT

The process of customer mall segmentation involves several steps. Firstly, random points are selected as centroids, and then the Euclidean distance of each point from randomly selected centroids is calculated. Based on the calculated distances, clusters are formed. The centroids are then re-calculated by finding the mean of the respective clusters obtained. This process is repeated until the constant value of centroids is obtained, and the latest cluster is considered the final cluster. The accuracy of the system is obtained as 0.9.

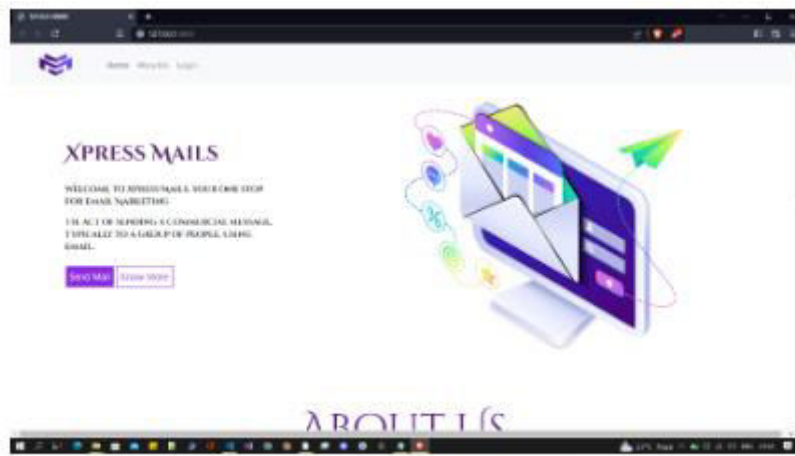


Fig.1: Home Page

The main objective of this project is to gather the mail dataset, followed by data pre-processing. After segmentation of the customer base using a machine learning algorithm, the segregated customers are then provided offers through email.

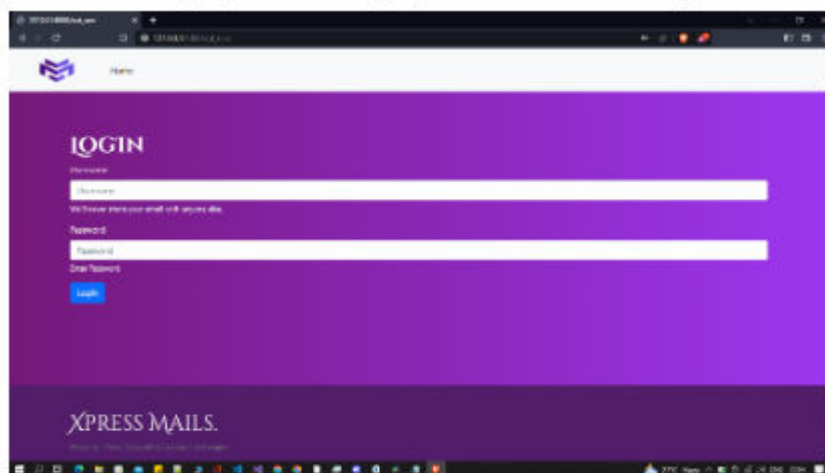


Fig.2: Login page

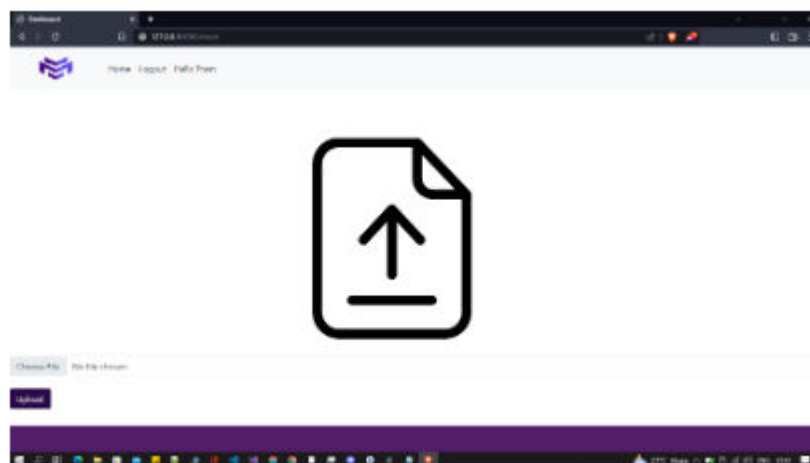


Fig.3: File uploading page

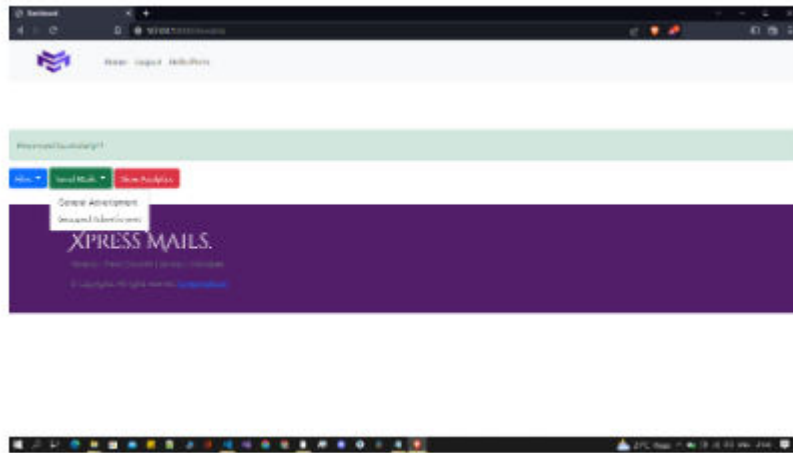


Fig.4: Main page

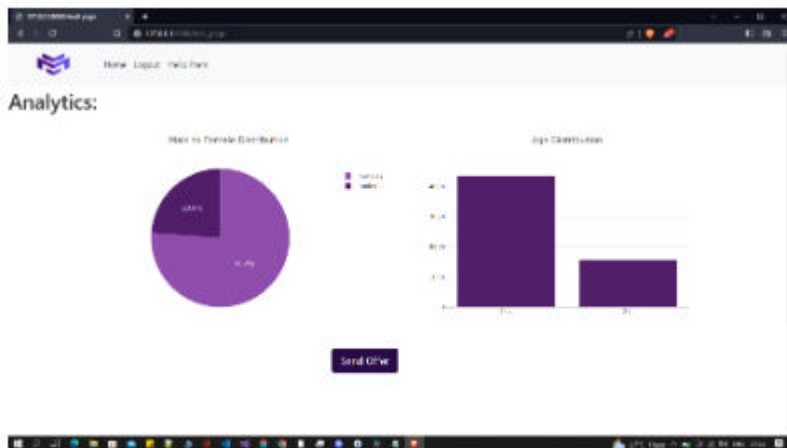


Fig.5: Analytics page

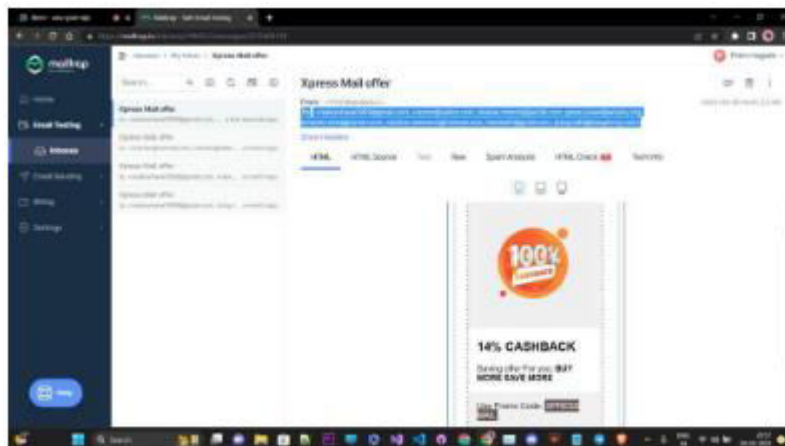


Fig.6: Email sending page

IV. CONCLUSION AND FUTURE SCOPE

Market segmentation is a practice that has arisen due to various factors such as the consumer movement, which has highlighted the underrepresentation of certain industries in the current system. As a result, managers are paying more attention to previously undisclosed client needs in response to these pressing demands. Another driver of market segmentation is intense competition, as companies seek to outperform their rivals by identifying untapped market niches. Additionally, the increasing use of technology for non-commercial purposes such as politics, religion, and public concerns has led to its frequent use in marketing efforts. Based on these factors, it can be concluded that:

Targeting the highest earning, highest spending customers can be an effective strategy, as they have more disposable income and can spend freely. Attracting high-income, low-spending customers can be achieved by soliciting feedback and improving product advertising. The average income and expenditure of customers who fall into the middle category may or may not benefit mall owners. To attract low-income, high-spending customers, businesses can offer incentives such as affordable EMIs. Low-income, low-spending customers may not be targeted by these businesses since they have modest incomes and make infrequent purchase.

REFERENCES

- [1] Customer Segmentation using K-Means Algorithm D P Kushwaha, D. P. Yash Kushwaha, "Customer Segmentation using K-Means Algorithm," 8th Semester Student of B. tech in Computer Science and Engineering.
- [2] Jan 2016 A R Ansari, A. R. Azarnoush Ansari, ""Customer Clustering Using a k-mean algorithm" in International Journal of Business and Management, Volume 11, Issue 7, 2016.
- [3] A Survey on Different Clustering Algorithms in Data Mining Technique Feb 20132249-6645 P. Indirapriya, Dr D K Ghosh, P. IndiraPriya, Dr. D.K. Ghosh " A Survey on Different Clustering Algorithms in Data Mining Technique" of International Journal of Modern Engineering Research (IJMER), Vol.3, Issue.1, Jan-Feb. 2013 pp-267-274 ISSN: 2249-6645.
- [4] CUSTOMER SEGMENTATION USING MACHINE LEARNING AMAN BANDUNI, Prof ILAVENDHAN School of Computing Science & Engineering, Galgotias University, Greater Noida, U.P.



INNO  **SPACE**
SJIF Scientific Journal Impact Factor
Impact Factor: 8.379



ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 **9940 572 462**  **6381 907 438**  **ijircce@gmail.com**



www.ijircce.com

Scan to save the contact details