International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Customer Segmentation Using RFM Model Machine Learning

Sanika Vikas Kamble

AISSMS Institute of Information Technology Near RTO, Pune Station, Pune-411001(MH), INDIA

A B S T R A C T

In the business world, it is crucial to know the client's desires and behavior patterns concerning buying merchandise. With the giant number of merchandise the businesses square measure confused to work out the potential customers to sell their merchandise. To solve this real-time downside we tend to use machine learning techniques and algorithms. A computer algorithm called the kmeans clustering algorithm is employed for this classification. The corporation can target specific clients with the help of formed clusters and market the material to them via social media and marketing campaigns. These strategies are effective when behavioral factors are taken into account.

Keywords: Machine learning, Customer segmentation, K-means algorithm, RFM model,

1. Introduction

The goal of many businesses is to stand out in the market by offering the best product and top-notch service. But they also need to watch out for slipping out and looking for alternatives. This is because they are the reason the company exists and because they develop their own market share.

Each consumer is unique in that they vary in terms of their age, region, and psychological makeup, but the most important factor is their purchase history. buying habits Every consumer category, therefore, needs a distinct offering at different pricing. In order to create a unique marketing plan for each group of customers, we must first classify related customers into a single segment. This essay's purpose is to determine the kind of customer(super customer, intermediate customer, basic customer and to estimate the worth of the customer to help businesses determine which kind of customer brings in substantial revenue, which does not, and what innovation marketing technique they can apply to increase their income growth by applying a new strategy.

This paper is to segment customers based on transactions from a supermarket which include 200 data entries. Here the clustering was done to help the retail industry to develop a new market strategy [4]. This paper user a Hierarchical clustering algorithm that does not prerequire information about the of clusters required.

2. Literature Survey

There are some previous studies related to the segmentation of customers some of them are listed below:

- Recent Trends in Intensive Computing M. Rajesh et al. (Eds.) © 2021 The authors and IOS Press. This article is published online with Open Access by IOS Press and distributed under the terms of the Creative Commons Attribution Non-Commercial License 4.0 (CC BY-NC 4.0). doi:10.3233/APC210200. Customer segmentation is performed on the company's customers' data and with the help of the K-means clustering machine learning algorithm customers are divided using features like total spending and annual income, this study also proves that dividing customers on the basis of behavioral characteristics is a better solution for existing customer segmentation problem and K-means clustering algorithm is identified as a good choice for this approach
- Assistant Professor, Computer Science and Engineering, Qis College of Engineering and Technology, Ongole, Andhra Pradesh, India, "Customer Segmentation Analysis and Visualization", International Journal of Science & Engineering Development Research Volume 7, Issue 1 Page Number: 280-284, January-February-2021 The algorithm used here is k-means algorithm A code is developed in python and it's trained on an information set having 201 data samples that are taken from the native shopping center.

Features and methods-

Stages of business focus on understanding the purpose of needs based on business valuation. After understanding the business initial data mining plan is designed to reach the goal. The study of this paper is of an online retail E-commerce website of a UK retailer. The transaction is for the years 2010 and 2011. here they have made clusters of customers using k –means clustering using the RFM model where customers are classified as super customers, and

intermediate customers based on their purchase history. Here they have used the RFM model(Recency Frequency Monetary): Recency, frequency, and monetary values are analysis tools being used to identify any organization's best customers.

Methodology used :

- Business understanding
- Data understanding
- Data Preparation
- Modelling
- Evaluation

Since it is straightforward and has advantages over other algorithms, the K-means technique is used for clustering in this study work.

K-Means algorithm

The most well-known and often employed technique for classifying data into groups and determining the appropriate number of clusters is K-means.

K-means is an iterative algorithm that attempts to divide the data into k different groups. The Elbow approach, which we shall explore later in the study, is used to predict the number of clusters that will develop in this situation, K.

Steps to take while utilising the K-means algorithm for clustering• Predict the number of clusters K.

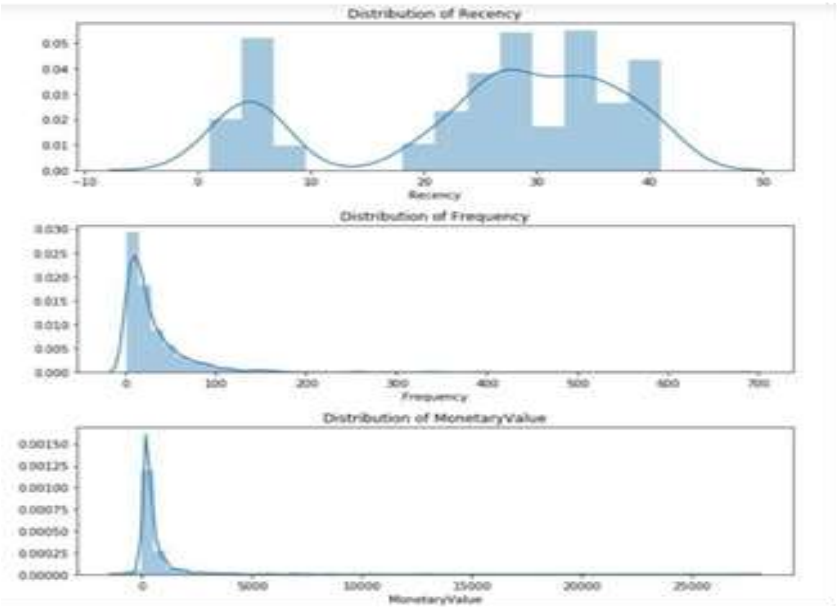
- Choose K data points at random to initialise the Centroid.
- Use all of the centroids to calculate the distance between the following data points.

The closest cluster should receive the data point; do this again.

Formula for Centroid Determination: Ci = 1 /M ∑j=1 m Xj Formula for Euclidean Distance d (p, q) = √(p1-q1)2 + (p2-q2)2 Elbow Method Elbow method is used to determine the optimal number of clusters based on the dataset. The idea is simple behind it, i.e., plotting the SSE (Sum squared Error) against suitable no of cluster value

Evaluation

The Silhouette Index is a number that is used to assess the consistency of the interpretation and validation within groups of data. This approach/technique provides a succinct graphical representation of the accuracy with which each object in a dataset is categorized. The object's similarity to its own cluster as opposed to other clusters is measured by the silhouette index. Silhouette value runs from -1 to +1; a high number denotes a good match between the item and the silhouette, and vice versa. The Silhouette Index may be used to choose the appropriate cluster design; for example, if many points are low or negative, the clustering arrangement may have a lot or little in the way of clusters[1]. S(i) = bi-ai / max (bi, ai) is the formula, where bi is the average distance between i and the cluster. ai = The typical separation between i in various clusters.



CONCLUSION:

- Customer segmentation is performed on the transaction data and with the help of K-means clustering machine learning algorithm customers are divided using different classes , this study also proves that the dividing customers on the basis of behavioral characteristics is a better solution for existing customer segmentation problem and K-means clustering algorithm is identified as a good choice for this approach.
- Based on the process of making a customer segmentation which is based on RFM model using Kmeans algorithm on a transaction data of a UK online retail store, we categorised the customer into 4 clusters based on the characteristics. These 4 clusters are basically Class A, Class B, Class C and Class D. Where Class A generates the highest revenue and Class D least. Customer segmentation is a very powerful tool to get the business insights and on how the customer behave. The value of silhouette index is 0.442 which is considered as good for the given dataset. Based on this result we obtained, this can help the company to develop market strategies and also can use as a promotional medium to their loyal customers. There are many other tools and method which can be used as comparison to the system that are already developed.

REFERENCES

[1]. Assistant Professor, Computer Science and Engineering, Qis College of Engineering and Technology, Ongole, Andhra Pradesh, India, "Customer Segmentation Analysis and Visualization", International Journal of Science & Engineering Development Research Volume 7, Issue 1 Page Number: 280-284 ,January-February-2021

[2]. Customer Segmentation Using Machine Learning Prof. Nikhil Patankar a ,1, Soham Dixit a , Akshay Bhamare a, Ashutosh Darpel a and Ritik Raina a a Dept. Of Information Technology Sanjivani College of Engineering, Kopargaon423601 (MH), India 2021

[3]. Rahul Shirole, Laxmiputra Salokhe, Saraswati Jadhav Department of Computer Engineering, Vishwakarma Institute of Technology Pune, Maharashtra, India Volume 8, Issue 3 Page Number : 591-597 May-June-2021