JETIR.ORG ISSN: 2349-5162 | ESTD Year : 2014 | Monthly Issue
JOURNAL OF EMERGING TECHNOLOGIES AND INNOVATIVE RESEARCH (JETIR)
An International Scholarly Open Access, Peer-reviewed, Refereed Journal

CUSTOMER SEGMENTATION USING MACHINE LEARNING

[1] [2] [3] [4]
Mrs.P.Rajeshwari, G. Shravya , Jeet S Patel , K.Dinesh

[1] *Assistant professor Department of Computer Science and Engineering*

Anurag University, Hyderabad, Telangana, India

[2,3,4]
Department of Computer Science and Engineering

Anurag Group of Institutions, Hyderabad, Telangana, India

ABSTRACT:

Customer Segmentation is the manner of dividing customers who might be applicable to advertising based on criterias such as gender, age, interests, and miscellaneous spending habits. The main purpose for any organization and enterprise is to recognize their centered clients, How their purchasers perform and use their offerings. Every customer may also use a organization's offerings differently. The trouble we're seeking to remedy is to outline this organization's purchasers. To outline positive behaviours and techniques those purchasers use the organization's offerings for. Companies that set up customer segmentation are in a belief that each client has special necessities and require a selected advertising attempt to cope with them appropriately. Companies intend to benefit a deeper expertise of the client they are targeting. Therefore, their intention needs to be unique and ought to be tailor-made to cope with the necessities of each and every client. Furthermore, through the information collected, groups can benefit a deeper expertise of client alternatives in addition to the necessities for coming across treasured segments that might achieve them most profit. This way, they could strategize their advertising strategies correctly. Customer Segmentation is one of the common application of unsupervised machine learning.. In this paper we have proposed a solution based on K-Means clustering which is a powerful method for clustering unlabelled dataset. The technique is to discover key attributes using which the customers can be grouped and gather few insights after visualizing the information. Data associated with demographics, geography, economic popularity in addition to behavioural styles can play a critical role in figuring out the organizations purpose in identifying their customer base.

Keywords – Cluster , K-means , Elbow , Exploratory Data Analysis.

1. INTRODUCTION

Here the aim is to find out the clusters of customers and therefore find out appropriate measures to increase income for the business. For example, a group of customers have immoderate income but their spending score (amount spent inside the mall) is low, so from the rating we are able to convert such form of customers into functionality customers (whose spending score is immoderate) through the manner of using strategies like better marketing, accepting remarks and improving our products. To find out such customers, the proposed method finds clusters based on different criteria.

1.1 MOTIVATION

Customer Segmentation permits in identifying least and most profitable customers, therefore helping the businesses to focus marketing activities on those customers who are most probably to buy your products or services which can help increase profits of the company.

1.2 PROBLEM DEFINITION

Customer Segmentation is a great example of unsupervised learning. Using clustering, we find out segments of customers by using the dataset of customers. Our method divides customers into various groups based on characteristics such as gender, age, interests,and spend score.

1.3 OBJECTIVE OF THE PROJECT

Nowadays the competition among companies is huge, thus every company aims to increase their profits through one way or the other. For every business corporation the most crucial issue is statistics. With the help of grouped or ungrouped statistics, we are able to perform some operations to find out consumer interests. Data mining is useful to extract statistics from the database in a human readable format. But, we may not apprehend the actual beneficiaries withinside the entire dataset. Customer segmentation is based mostly on consumer's age, demographics, spend score, income, gender, etc. By this, we are able to apprehend inferences such as which products give huge profits to the company or which age of people do the most shopping etc. Thus company focuses mostly on those ages of people and those products which have a greater potential to generate maximum revenue for the company.

2.LITERATURE SURVEY

A. CUSTOMER CLASSIFICATION

Over the years, the commercial world has become extremely competitive, as companies try to satisfy their existing customers through their products or service, attract new customers, enhancing their businesses. [6] The job of working out and identifying the desires and the needs of every consumer is an extremely tidious task. The reason is consumer's desires differ according to their wants, needs, geographic location, size, taste and product features. Further, it turns out that treating each consumer equally is a bad exercise. This project has followed the idea of consumer segmentation or marketplace segmentation, whereby shoppers are divided into subgroups or segments, whereby contributors of each subcategory inhibit similar behaviours. Apparently, customer segmentation is the means of dividing market into different areas of target.

B. Data Repository

Data gathering or Data collection is still under research in areas such as physical and social sciences, enterprise. The motive of collecting data is to reap quality proof that leads the analysis to build concrete solutions to the queries given.

C. Clustering Data

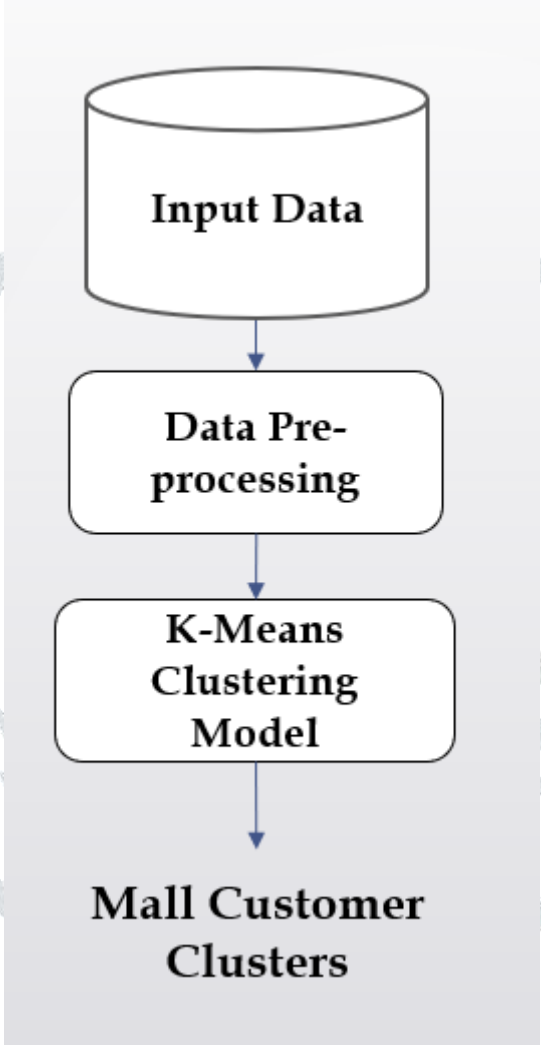
Clustering is the manner of grouping data based on some features or characteristics. There are various algorithms, which can be enforced on datasets based on the supplied condition.[7] But, there exists no certain algorithm which can be applied for all clustering problems, thus it becomes crucial to choose out the appropriate clustering techniques.

D. K-Means

K-means algorithm is one of the commonly used clustering algorithm. This clustering algorithm uses the concept of centroid, whereby each datapoint is placed in one among the overlapping ones, which are already taken care by the K-Means algorithm. The mentioned method uses elbow technique to determine the K value.

3. METHODOLOGY

Firstly we explored the dataset and obtained some insights which can be useful for our model. Next we cleaned the dataset by handling nan values or NULL values, removed duplicate records etc. Then we used visualization tools to visualize our data and obtain important information from our data. Data preprocessing was done to make the data ready to fit the model which included feature scaling. Before implementing the K-means algorithm our Finally the K-means clustering model was built and the output clusters were displayed using a scatter plot.



4. CONCLUSION

Nowadays the competition has been highly increased in every industry, retail is no exception. So every business either it may be a small supermarket or an ecommerce giant like amazon ,flipkart. Every business try to use some tools, approaches, marketing strategies to attract customers towards their business. One such approach used by the above mentioned is customer segmentation. It is obvious that each and every customer can't be served with same product model, SMS campaigns, emails, advertisements. Customers have different needs. Treating all customers equally might not benefit the company in long run. Customer segmentation is one such cure for this problem. Finding optimal number of unique customer groups will help you understand how your customers are different, and help you give them exactly what they expect from your company. Employing Customer segmentation has high probability of increasing your company's revenue. This is the reason why segmentation can turn out to be a great technique by means you can surpass your competitors in terms of profits and can get you more customers. Doing it with machine learning is definitely the right choice.

5. FUTURE ENHANCEMENTS

While this method proposes a step-by-step manner for identifying and focusing on your best customer segments, truly following it does not completely gurantee company's profits.. To be efficient, you have to put together and plan for the diverse demanding situations and hurdles that may occur at every step, and continually make certain changes to the method to process any new incoming data that would alternate its output. And only employing this method cannot produce expected results, this method can also be collaborated with some other techniques, tools to produce best possible results.

6. REFERENCES

[1] I. S. Dhillon and D. M. Modha, "Concept decompositions for huge sparse textual content records using clustering," Machine Learning, vol. 42, trouble 1, pp. 143-175, 2001.

[2] T. Kanungo, D. M. Mount, N. S. Netanyahu, C. D. Piatko, R. Silverman, and A. Y. Wu, "A Green K-manner clustering algorithm", IEEE Trans.Pattern Analysis et Machine Intelligence, vol. 24, p. 881892, 2002.

[3] D. Aloise, A. Deshpande, P. Hansen and P.Popat, "Basics of market segmentation"

Euclidean sum-of-squares clustering," Machine Learning, vol. 75, pp. 245-249, 2009.

[4]Puwanenthiren Premkanth, —Market Segmentation and Its Impact on Customer

Satisfaction with Especial Reference to Commercial Bank of Ceylon PLC.‖ Global Journal of Management and Business Research Editeur : Global Journals Inc. (USA). 2012. Stampa ISSN : 09755853. Volume 12 Edizione1.

[5]Irene Gil Saura and Maria Eugenia Ruiz Molina, "Customer Segmentation Based on Engagement and Use of ICT", Industrial Management & amp; Data Systems, vol. 109, no. 2, p. 206223, 2009.

[6] Stone and Bob, Successful Direct Marketing Methods[M]. 4 ed. NTC Business Books. 1989:10̃110

[7] Management Science: A Comparative Research on the Methods of Customer Segmentation Based on Consumption Behavior. 2003.2, Vol.16.Google Scholar

[8] Jayant Tikmani, Sudhanshu Tiwari and Sujata Khedkar, "Mall customer segmentation based on cluster analysis An Approach to Customer Classification using k-means", IJIRCCE, 2015.

[9] Jing Wu et Zheng Lin, « Research on the customer segmentation model by clustering », ICEC, 2005.