

CUSTOMER SEGMENTATION

¹Riya Kanchan, Shubhangi Sangal, Prof Upendra Tiwari

Student, ABES Institute of Technology, Student, ABES Institute of Technology, Assistant Professor, ABES Institute of Technology
Computer Science and Engineering,
ABES Institute of Technology, Ghaziabad, India

Abstract : Ecstatic and joyous customer is commensurate to the company's growth. More the level of satisfaction the more is the probability of the increase in the company's net worth. Customer segmentation is done to fathom the behavior of customer and grasp information regarding his interest and need for a particular product. Customer segmentation has made world more customer centric than product centric. Various data mining techniques had been used previously for the customer classification. In this paper we will be discussing about the use and advantage of k-means clustering data mining algorithm over other data mining techniques. Based on the essence of demographic, geographic, behavioral, and psychographic conditions of the customer stratification will be done. We will be doing this using by first classification and then applying k means clustering algorithm in weka tool and we will be able to decide the best among them and suggest what technique should be taken up in order to do the segmentation of the customer.

IndexTerms - customer segmentation, data mining, k-means clustering, weka tool .

INTRODUCTION

Customer segmentation is the effective way of arriving at a smart decision that which attribute contribute most to the buying pattern of the customer. Customer of different age, gender, location, marital status etc have different requirement and outlook towards a product. Customer are people of utmost importance for an organization or company. The business cannot flourish and has least scope of growth without the support of satisfied customer. So, in order to attain victory every organization and firm focus at grasping the conduct and qualities which are possessed by the particular customer.

Effective customer segmentation not only helps the company to group information needed for classification but in return it also up to some extent guarantee customer's fidelity concerning the firm. The more precisely the classification is done more is the level of satisfaction and more is the chances of customer to prefer the company over others. However, it is challenging task to categorize the type of customer, his/her need, his/her buying pattern etc.

Company or organization works on different attributes and character of customer but is not known to which factor they pay more importance. Various approaches are being used for the customer segmentation. If we broadly classify them they are demographic, geographic, behavioral and psychographic. Geographic factor signifies the location, state, country and its density and size where the customer resides. Demographic component signifies the maturity, sex, bachelorhood or spousal relationship, number of children, education, earnings, religion and nationality of the customer. Behavioral component signifies the event on which a customer purchases a product, fidelity, outlook, advantage which customer takes into consideration. We will be considering and discussing geographic and demographic factors in this paper.

Various mining techniques had been used previously to deduce the result. Data mining is a process in which we observe the information stored in the database and retrieve important information and obtain pattern for it. It is a strong and vital tool which is helpful for the organization and company as it help them to predict the future trend and the changing customer demand and classify customer on that basis. Business doubt and conclusion which are laborious to make can be solved using data mining. K-means clustering algorithm is one of the data mining technique which can be used to categorize the customer.

Literature Survey

Clustering

Clustering can be described as the practice of assembling together a suite of data in such a manner that data in similar cluster are more analogous to each other than data in other cluster. Extensively we can split it into two parts. Hard and soft are the two known types.

- Hard clustering-In this type single data belongs to sole and only sole cluster.
- Soft clustering-The alternate name for it is fuzzy clustering. In such type of clustering, it is noticed that the sole data can be owned by more than one cluster.

Data Mining

In this practice a huge aggregate of initially available data is scrutinized in an endeavor to attain new beneficial facts from it. Types of Data Mining Algorithms

1. Apriori-This approach of algorithm is implemented on the database having a large amount of data and depends upon association rules. These rules and regulations assist in knowing out the connection that exist in the middle of the data items accommodated in the database.
2. Cart-Cart stands for Classification and Regression trees. In such type binary trees are constructed.
3. C4.5-C4.5 constructs the classifier in form of a decision tree. The data we want to classify is done with the help of classifier tool present in data mining which help in predicting that data belong to which category.
4. K-means-In this algorithm we split the n-observations into K-cluster where the data in every cluster is linked to one another and unlike from that in the other cluster.

Tools used for clustering

1. SPSS-SPSS is a tool of data management and statistical analysis which can be utilized for broad domain of data. With the assistance of foregoing trends and precise forecast it aid in constructing efficient model.
2. ANOVA- It assist in suggesting that whether there prevail a difference between the group or not.
3. IBM Data Miner- IBM Data Miner is the data mining tool of IBM.
4. Weka- It incorporates a collection of algorithm of machine learning for solving issues that our encountered in data mining. The language utilized for learning it is java language and any platform can be used to execute it. We can either instantly apply these algorithm to dataset we own or we can call it from java code of our own.

Advantages of weka tool

- It is helpful in developing new machine learning schemes.
- Since it is open source, it has free availability.
- Its main advantage is that since it operates on java it can run on any platform.
- It is easy to use due to its graphical user interface.

Methodology

We have designed an online billing system through which we entered the data and created the bill of customer. The data get saved into database from where after converted into csv format we loaded it into weka and found the best segmentation.

- Collection of data.
- Conversion of database data into csv format.
- Loading of data into weka.
- Apply classification and clustering on data.
- Find out the best way to segment data from result.

The tools and technology used here are

- Weka tool.
- Editor: Notepad
- JSP
- HTML
- MySql
- Web logic 12c

Conclusion

The results we obtained after performing the segmentation on customer's data by K-means, Apriori and classification algorithm showed us that the Behavioral segmentation is the best type of customer segmentation for segmenting the customer's data among Geographic customer segmentation, Demographic customer segmentation and Psychographic customer segmentation and Weka tool is the efficient tool for performing this type of clustering. Customer segmentation based on the Behavioral type helps the customer or the user and the organization or company both to yield the best possible result possible.

Acknowledgment

I want to give special thanks to all my teacher, supervisor and guide of my project, Department of computer science and engineering, Abdul Kalam Technical University for their continuous support and guidance. I highly obliged for entire support to complete my project successfully. Finally, I express my real thanks to my family, friend and my project supervisor who always inspired and motivates me to complete my project propitiously during my project work.

REFERENCES

- [1]. Prof.Tejal Upadhyay, Atma Vidhani, Vishal Dadhich," Customer Profiling and Segmentation using Data Mining Techniques", International journal of computer science & communication, Vol.7,2017.
- [2]. Kishana R. Kashwan," Customer Segmentation Using Clustering and Data Mining Techniques", International Journal of Computer Theory and Engineering, Vol. 5, No. 6, December 2013.
- [3]. Dr. Sankar Rajagopal," customer data clustering using data mining technique", International Journal of Database Management Systems, Vol.3, No.4, November 2011.
- [4]. Hasan Ziafat, Majid Shakeri," Using Data Mining Techniques in Customer Segmentation", Int.Journal of Engineering Research and Applications, Vol. 4, Issue 9(Version 3), September 2014.
- [5]. Mythili, Madhiya, "An Analysis on Clustering Algorithms in Data Mining", International Journal of Computer Science and Mobile Computing, Vol.3 Issue.1, January- 2014.