iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 5 **Issue:** II **Month of publication:** February 2017

DOI: http://doi.org/10.22214/ijraset.2017.2033

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)

Analysis of Clustering Technique in Marketing Sector

Pallavi R. Wankhade[1], Prof. Rajeshri R. Shelke[2]
[1]ME Second Yr (CSE) H.V.P.M's COET, Amravati
[2]Assistant Professor,ME,Ph.D(Pursuing) H.V.P.M's COET, Amravati

Abstract: ***Cluster analysis divides data into meaningful or useful groups (clusters). One of the most important problems in modern finance is finding efficient ways to summarize and visualize the stock market data to give individuals or institutions useful information about the market behaviour for investment decisions. The enormous amount of valuable data generated by the stock market has attracted researchers to explore this problem domain using different methodologies. Potential significant benefits of solving these problems motivated extensive research for years. We proposed clustering techniques that are being used in Data Mining is presented. Data mining adds to clustering the complications of very large datasets with very many attributes of different types. This imposes unique computational requirements on relevant clustering algorithms with k-means method is one of the clustering techniques. Data mining facilitates marketing sector by classifying customer demographic that can be used to predict which customer will respond to a mailing or buy a particular product and it is very much helpful in growth of business. If meaningful clusters are the goal, then the resulting clusters should capture the "natural" structure of the data. For example, cluster analysis has been used to group related documents for browsing, to find genes.***
Keywords: Marketing, data mining, clustering technique.

I. INTRODUCTION

Clustering is the process of grouping a collection of objects into classes of similar objects. Cluster analysis is a very important tool in data analysis. It is a set of methodologies for automatic classification of a collection of patterns into clusters based on similarity. Cluster analysis has wide applications in data mining, information retrieval, biology, medicine, marketing, and image segmentation. With the help of clustering algorithms, a user is able to understand natural clusters or structures underlying a data set. For example, clustering can help marketers discover distinct groups and characterize customer groups based on purchasing patterns in business.

Advancements in computer technologies caused a rise in information production and data base system volume. To discover the data with the potential to be useful which are kept in databases and to create meaningful patterns from these are stated as data mining [2]. Businesses are in a tense competition which needs continuity in today's consumer focused markets. Businesses have to apply effective and low cost marketing strategies to be successful in these competition conditions to create effective marketing strategies true information is needed and to obtain true information future headed forecasting systems which can analyze the data in multiple dimensions are needed. In this connection, the data mining techniques are used widely in marketing field same as many other fields.

For this method of clustering we start by deciding how many clusters we would like to form from our data . In the K-means algorithm, K value can be determined according to problem or it can not be determined. Like squared error criterion, there is need to have a clustering criterion. The K-means algorithm starts with random selection of an object that will represent every cluster. Each of remaining objects is assigned to a cluster and the clustering criterion is used to compute average of the cluster.

II. LITERATURE REVIEW

Radhakrishnan et al.,[1] described the Data Mining Application in Marketing .Data mining is applied to this problem by first defining what it means to be a good prospect and then finding rules that allow people with those characteristics to be targeted. Prospecting requires communication. Broadly speaking companies intentionally communicate with prospects in several ways. Partitioning method is referred as centroid based clustering such as K-means and partitioning around mediods. The clustering technique also plays a significant role in data analysis and data mining applications [3,5,6,8].

Chandan et al.,[7] are developed A new Approach Document clustering help in organising documents in groups according to their similarity of contents.

Zhang et al.,[9] Marketing is one of the mostly used application area for Data mining by the industry in general. Banking is not an

©IJRASET: All Rights are Reserved

exception. Retaining customers and finding new customers are getting increasingly difficult because of cut throat competition prevailing in the market these days. This is where data mining can help a great deal.

In traditional markets, customer clustering is one of the most significant methods used in studies of marketing. This study classifies existing customer cluster/segmentation methods into methodology-oriented and application-oriented approaches. Most methodology driven studies used mathematical methodologies; e.g. statistics, neural net, generic algorithm (GA) and Fuzzy set to identify the optimized segmented homogenous group [2,4].

Chopra et al.,[10] Data mining applied to customer relationship management systems can analyze customer data. Data mining techniques can help in classifying customers according to the customer's, attributes, behaviour, needs, preferences, value and other factors. Sreekumar et al.,[3] determined how customers will react to a change in interest rates, which customers will be likely to accept new product offers, what collateral would require from a specific customer segment for reducing loan losses.

III. EXISTING SYSTEM

In traditional markets, customer clustering is one of the most significant methods used in studies of marketing sector. This study classifies existing customer cluster/segmentation methods into methodology-oriented and application-oriented approaches. Most methodology driven studies used mathematical methodologies; e.g. statistics, neural net, generic algorithm (GA) and Fuzzy set to identify the optimized segmented homogenous group .Retaining customers and finding new customers are getting increasingly so it difficult to identify. One of the most important problems in modern finance is finding efficient ways to summarize and visualize the stock market data to give individuals or institutions useful information about the market behaviour for investment decisions. The enormous amount of valuable data generated by the stock market has attracted researchers to explore this problem domain using different methodologies.

A. Categorization of Clustering Techniques and Previous Work

Many diverse techniques have appeared in order to discover cohesive groups in large datasets. In the following two section we present the two classic techniques for clustering as well as more specific ones, respectively. A large number of clustering algorithms have been developed for different purposes. Based on the strategy of how data objects are distinguished, clustering techniques can be broadly divided in two classes: hierarchical clustering techniques and partitioning clustering techniques. However there is no clear boundary between these two classes.

1) *Partitioning methods:* Partition clustering algorithms divide the data set into a specified number of clusters. Partitioning clustering algorithms, such as K-means, PAM (Partition around Mediods) or K-mediods and CLARANS (Clustering Large Application Based on Randomized Search). K-means is the most popular and easy-to understand clustering algorithm. The K-means Algorithm most widely used partitional algorithm is the iterative K-means approach. The algorithm is called *k*-means due to the fact that the letter *k* represents the number of clusters chosen.
2) *Hierarchical methods:* Hierarchical clustering, also known as Connectivity based clustering, is based on the core concept of objects being more related to nearby objects than to objects farther away. Hierarchical clustering is a method of cluster analysis that constructs the clusters or groups by recursively partitioning the instances in either a top-down or bottom-up approach. Hierarchical clustering algorithm builds a cluster hierarchy or a tree of clusters.

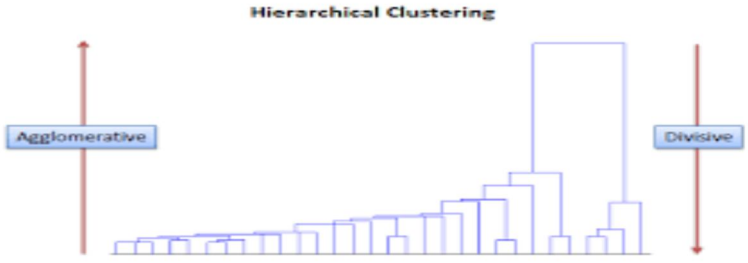


Figure 1: Hierarchical Clustering

3) *Agglomerative Method:* It is a bottom-up approach which starts by assigning each data instance to one cluster and then

©IJRASET: All Rights are Reserved

iteratively merges the two most similar clusters. This technique builds the hierarchy from the individual objects by progressively merging clusters which is shown in figure 1.

4) *Divisive Method:* t is top-down approach which starts by assigning all the objects to one cluster and then iteratively dividing it into smaller and smaller clusters which is shown in figure 1.

B. Application

Clustering techniques are useful in various applications of real world including data/text mining, voice mining, image processing, web mining etc. It is a main task of exploratory data mining, and a common technique for statistical data analysis used in many fields, including machine learning, pattern recognition, image analysis, information retrieval, and bioinformatics .

1) *Application Areas of Clustering:* Clustering algorithms can be applied in many fields, for instance

Marketing: finding groups of customers with similar interests and behaviour given a large database of customer data containing their properties and past buying records. Medicine: IMRT segmentation, Analysis of antimicrobial activity, Medical imaging.

Financial task: Forecasting stock market, currency exchange rate, bank bankruptcies, understanding and managing financial risk, trading futures, credit rating. Computer Science: Software evolution, Image segmentation, Anomaly detection. Biology: classification of plants and animals given their features, human genetic clustering, transcriptomics. Insurance: identifying groups of motor insurance policy holders with a high average claim cost; identifying frauds. ity-planning: identifying groups of houses according to their house type, value and geographical location. Earthquake studies: clustering observed earthquake epicentres to identify dangerous zones.WWW: document classification; clustering web log data to discover groups of similar access patterns.

2) *Advantages:* It does not require a-priori specification of number of clusters which is opposite from the case of k-means.

is able to identify noise data while clustering, so more robust in nature. Able to find arbitrarily size and arbitrarily shaped clusters.

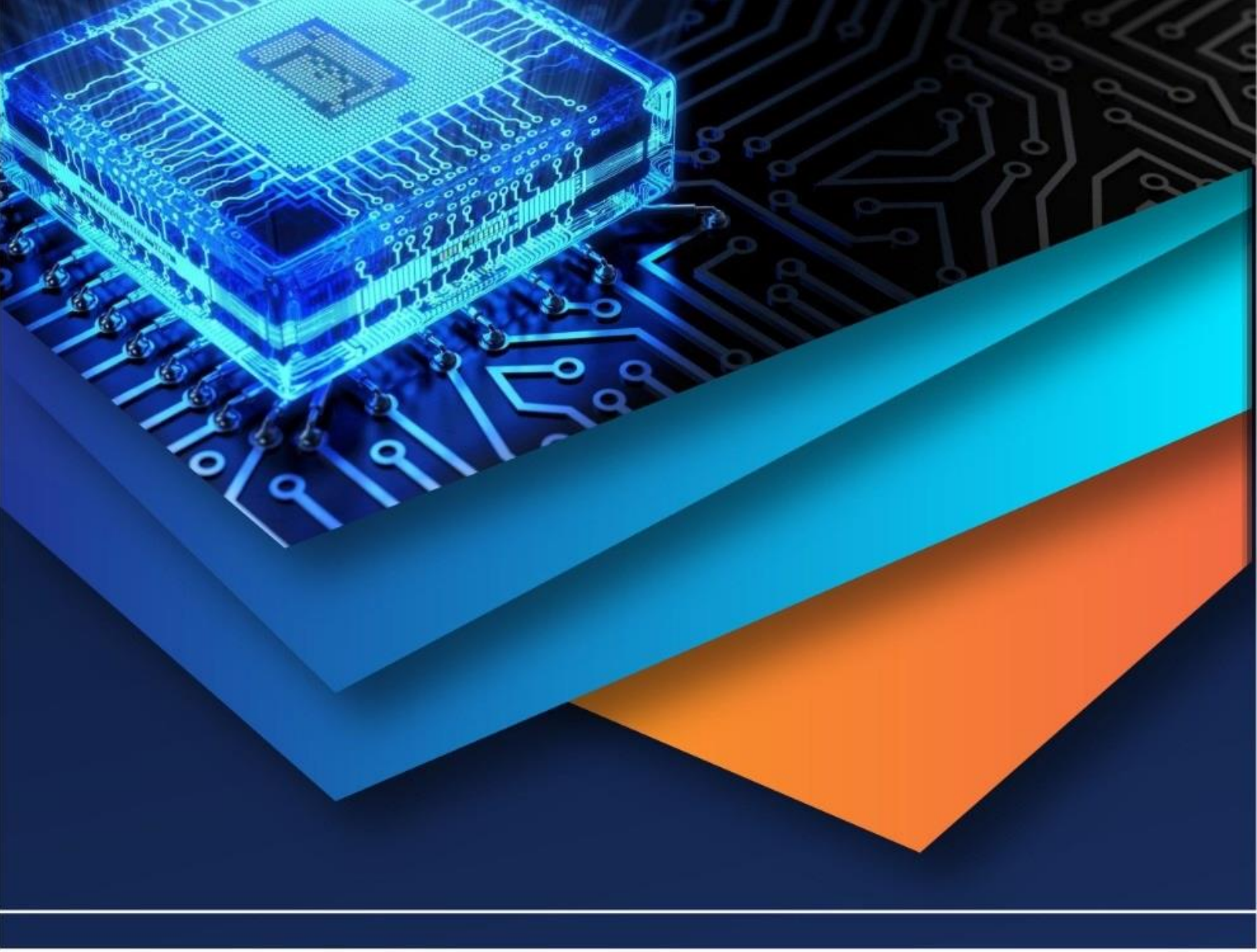
IV. CONCLUSION

With the increase of economic globalization and evolution of information technology, financial data are being generated and accumulated at an unprecedented pace. Data mining techniques will be used to uncover hidden patterns and predict future trends and behaviours in financial markets. The competitive advantages achieved by data mining include increased revenue, reduced cost, and much improved marketplace responsiveness and awareness. We described the process of clustering from the data mining point of view. We gave the properties of a “good” clustering technique and the methods used to find meaningful partitioning. These techniques are being used in marketing sector. This study concludes that clustering techniques and algorithms become a highly active research area in data mining research.

REFERENCES

[1] Radhakrishnan B, Shineraj G, Anver Muhammed K.M ,“Application of Data Mining In Marketing”, nternational Journal of Computer Science and Network, Volume 2, Issue 5, October 2013.
[2] Aslı Çaliş , Ahmet Boyaci , Kasım Baynal ,“Data mining application in banking sector with clustering and classification methods”,Proceedings of the 2015 International Conference on Industrial Engineering and Operations Management Dubai, United Arab Emirates (UAE), March 3 – 5, 2015.
[3] Sreekumar Pulakkazhy and R.V.S. Balan, “Data Mining In Banking And It’s Application A Review”, Journal of Computer Science 9 (10): 1252-1259, 2013.
[4] Dr. Sankar Rajagopal, “CUSTOMER DATA CLUSTERING USING DATAMINING TECHNIQUE”, International Journal of Database Management Systems (IJDMS) Vol.3, No.4, November 2011.
[5] Rajwant Kaur, Kiran Jyoti, Rohit Kumar “ A Review Paper on Knowledge Discovery and Data Mining Techniques”,International Journal of Advanced Research in Computer Science and Software Engineering ,Research Paper Available online at: www.ijarcsse.com,Volume 4, Issue 8, August 2014.Apurva
[6] Juyal, Dr. O. P. Gupta, “A Review on Clustering Techniques in Data Mining”, International Journal of Advanced Research in Computer Science and Software Engineering,Volume 4, Issue 7, July 2014.
[7] Chandan Jadon, Ajay Khunteta, “A New Approach of Document Clustering”, International Journal of Advanced Research in Computer Science and Software Engineering Research Paper,Volume 3, Issue 4, April 2013.S.R.Pande, Ms.S.S.Sambare, V.M.Thakre’ “Data Clustering Using Data Mining
[8] Techniques”, International Journal of Advanced Research in Computer and Communication EngineeringVol. 1, Issue 8, October 2012.
[9] R. R. Shelke ,Dr. V. M. Thakare, Dr. R . V. Dharaskar,“Study of Data Mining Approach for Mobile Computing Environment”,International Journal on Computer Science and Engineering (IJCSE) , ISSN : 0975-3397, Vol. 4, 12 Dec 2012 ,pp.1920-1923

©IJRASET: All Rights are Reserved



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)