CRM Application using Fuzzy Clustering Algorithm
Abhay Patil1 Utkarsha Mane2 Arati Kachare3 Madhura Mane4
1,2,3,4Rajiv Gandhi Institute of Technology, Mumbai Maharashtra, India
1,2,3,4Department of Information Technology

Abstract—Customer Relationship Management (CRM) is the process of managing a good relationship with customer and retain customers. The telecom industry provides the services to the customer in order to satisfy their needs and requirements. It handles the large data of customers, but due to market dynamity and competition, company face one problem that is loss of valuable customers. To overcome this, we need to understand the behavior of customers in terms of usage level of provided services. Data mining is the efficiently used tool in CRM. It collects all the information about the customers. Also extract the patterns from the data to transform into knowledge information. Clustering is one of the data mining techniques used in customer segmentation. This segmentation approach is achieve using fuzzy logic. With the help of this it forms the fuzzy cluster. This cluster represents the group of customers regarding particular services. This technique is referred to as soft clustering also called as fuzzy clustering. In Soft clustering each data point has a probability of being in each cluster. The data points are assigned memberships value for each of the clusters. The complete operation is known as Fuzzy C-means (FCM) Clustering algorithm. It also handles the extreme outliers by assigning them a very small membership degree in surrounding clusters. The formation of clustered data represents the output. This paper focuses on applying FCM algorithm to calculate the churn ratio of a particular service accurately.

Key words: Fuzzy Logic, FCM algorithm, Churn ratio

I. INTRODUCTION

In terms of the market dynamity and competition the telecom industry is rapidly growing. To satisfy customer needs and requirements it creates new technologies and products, which gives a series of options and offers. Telecom companies in particulars face the problem of a loss valuable customers. That’s why in general we can say that diverting the customers is always one of the most demanding issues that telecom industries tackles in normally. So solve these types of problems we have to understand the customers behavior and categorize them on the basis of their operation levels. So required decisions can be taken to intensify providing services. So data mining system, is effectively understand customer behavior and then it takes decisions.

Extracting patterns from data, which is called as the Data mining process. To transform data into business intelligence, it giving very effective advantage. So nowadays it is widely used in profiling services like marketing, surveillance, fraud detection and scientific discovery. Data mining process can help telecom companies to search effective modes, designs and connections in their customer relationships.

For any organization CRM is very important because it wanted to understand relationships between customer and their organization. CRM starts with intensive analysis of customer activities. Hence telecom industries need personal interactions with customers. So we must be congregate all customer related data. And after that analyzing this data, then we will achieve the customer gratification. Hence data mining is the model through which company can assemble all the information about customers.

The data mining model contains association rules, classification, clustering, regression analysis, sequence etc. In the customer segmentation clustering and classification. To identify minimize group dissimilarity within group based on a distance function. Hence telecom industry is based on only customer information from the usages of services. Which is provided from the telecom companies that explains customers and their operations with the company.

From the telecom industries customer data sets are very huge and are in terabyte levels. To handle those huge data in order to search result and anticipate customer fuzzy clustering techniques are the effective one. In this we analyze that how the algorithm implemented in large volume of real time telecom data sets. The result off algorithm shows the behavior of customers. From usage of services we can segregate services that high level, middle level and low level. And according to the analysis of output, segregation can be made between customer and after that CRM can be enhanced intensively in telecom industry.

II. BACKGROUND WORK

The telecom industry is rapidly growing in market. Company handles customer information in a large database. In the era of great competition, understanding and satisfying customer’s requirements are the critical tasks for the telecom company to make profits and retain the customers by handling huge datasets. So In this project our ultimate goal is to retaining the customers in Telecom Company.

CRM is about acquiring and retaining customers, improving customer loyalty, gaining customer insight, and implementing customer-focused strategies, building a relationship with each customer to improve customer satisfaction and maximize profits. It’s about understanding, anticipating and responding to customers’ needs. To manage the relationship with the customer a business needs to collect the right information about its customers and organize that information for proper analysis and action. The effective CRM package is not just in what data is collected but in the organizing and interpretation of that data. The process of finding useful patterns and information from raw data is often known as knowledge discovery in database. Data mining is a particular step in this process involving the application of specific patterns from data. Data mining is the efficiently used tool in CRM. Data Mining (sometimes called data or knowledge discovery) is the process of analyzing data from different perspectives and summarizing it into useful information - information that can be used to increase revenue, cuts costs, or both. Data mining software is one of a number of analytical tools for analyzing data. It allows users to analyze data from many dimensions or angles, categorize it, and summarize the relationships....
identified. Data mining helps to identify valuable patterns contained in diverse data and their relations so as to help the major decisions. Data mining can create the following models: (1) classification; (2) cluster analysis; and (3) association analysis [1].

In this paper, we apply the cluster analysis using Fuzzy C-means clustering algorithm. Cluster analysis is a technique for breaking data down into related components in such a way that patterns and order becomes visible. It aims at shifting through large volumes of data in order to reveal useful information in the form of new relationships, patterns, or clusters, for decision making by user. Clusters are natural groupings of data items based on similarity metric or probability density models. Clustering algorithm maps a new data items into one of several known clusters. In fact cluster analysis has the virtue of strengthening the exposure of patterns and behaviors as more and more data becomes available. A cluster has a center of gravity which is basically the weighted average of the cluster. Membership of a data item in a cluster can be determined by measuring the distance from each cluster center to the data point. The data item is added to a cluster for which this distance is a maximum.

In Telecom Company customer behavior is unpredictable in terms of usage level. It may vary day by day and their usage pattern is an intermediary, with the nature of “and also”, so in these sample uncertain situations, soft clustering called fuzzy clustering can be used efficiently and easily.

The central idea in fuzzy clustering is the non-unique partitioning of the data in a collection of clusters. The data points are assigned memberships value for each of the clusters. The fuzzy clustering algorithm allows the clusters to grow into their natural shapes. In some cases the membership value may be zero indicating that the data point is not a member of the cluster under consideration. Many crisp clustering techniques have difficulties in handling extreme outliers but fuzzy clustering algorithms tend to give them very small membership degree in surrounding clusters. The nonzero membership values, with a maximum of one, show the degree to which the data point represents a cluster. Thus fuzzy clustering provides a flexible and robust method for handling natural data with uncertainty. In fuzzy clustering, each data point will have an associated degree of membership for each cluster. The membership value is in the range zero to one and indicates the strength of its association in that cluster.

III. PROPOSED SYSTEM

A. Fuzzy C-Means:

In fuzzy clustering (also referred to as soft clustering), data elements can belong to more than one cluster, and associated with each element is a set of membership levels. These indicate the strength of the association between that data element and a particular cluster. Fuzzy clustering is a process of assigning these membership levels, and then using them to assign data elements to one or more clusters. Fuzzy clustering is an approach operating towards fuzzy logic and it provides the flexible method of assigns the data points to the clusters. In this technique membership value has been assigned to each cluster. Membership can also help us to decide whether the data points belong to the clusters or not.

One of the most widely used fuzzy clustering algorithms is the Fuzzy C-Means (FCM) Algorithm (Bezdek 1981). The level of cluster fuzziness is determined by the fuzzifier m. large is the m smaller is the memberships and hence, fuzzier Cluster. In the limit m=1, the memberships converge to 0 or 1, which is a crisp Partitioning.

I) Algorithmic steps for Fuzzy C-Means Clustering:

Let \( X = \{x_1, x_2, x_3, ..., x_n \} \) be the set of data points and \( V = \{v_1, v_2, v_3, ..., v_c \} \) be the set of centers.

1) Randomly select ‘c’ cluster centers.
2) Calculate the fuzzy membership ‘\( \mu_{ij} \)’ using:
\[
\mu_{ij} = \frac{1}{\sum_{k=1}^{c} \left( \frac{d_{ij}}{d_{ik}} \right)^{(2/m-1)}}
\]
3) Compute the fuzzy centers ‘\( v_j \)’ using:
\[
v_j = \frac{\left( \sum_{i=1}^{n} (\mu_{ij})^m \ x_i \right)}{\left( \sum_{i=1}^{n} (\mu_{ij})^m \right)}, \forall j = 1, 2, ..., c
\]
4) Repeat step 2) and 3) until the minimum ‘J’ value is achieved or \( \| U \( k+1 \) - U \( k \) \| < \beta \).

Where, ‘k’ is the iteration step.

‘\( \beta \)’ is the termination criterion between \([0, 1]\).

‘\( U \)’ = (\( U_{ij} \)) is the objective function, ‘m’ is the fuzzy membership matrix.

B. Fuzzy Clustering in Telecom Company:

Fuzzy clustering methods are great helpful to mine Telco customer and to find out their role in business market. To make the fuzzy clustering more impactful in traditional databases we have four key steps:

- Customers may have membership in more than one fuzzy cluster.
- The outcome of fuzzy clustering is not having any relation to the specific offer or market message.
- Clustering outcome differs with the various time offers and as the service changes .Such fuzzy clustering evolution helped to better analysis of loyalty of customers.
- For the longer time one can suggest formal procedures that involve intuitive fuzzy clustering for improvement.

IV. IMPLEMENTATION AND RESULTS

We have performed research to carry retention of customers in the telecom industry. We have collected the customer data. We are taking 7 services in our system. These 7 services are as follows:

- Failure of Calls
- Service_Complains
- Charge_amount
- Seconds_of_use
- Freq_of_use
- Freq_of_sms
- no_of_distinctCalls

Hence on our selected 7 services and 90 customer’s data we have applied Fuzzy C-means i.e. Soft partitioning algorithm to accurately find the churn rate of customers in telecom industry.
The implemented result is shown below:

Input given:
No. of customers=100, No. of services=7, Algorithm used=Fuzzy C-means

A. Result:

FCM performs clustering on the basis of membership value of customers. This membership value gives the accurate churn ratio of services used by the customers moving from one cluster to other cluster. The result of soft clustering is more accurate because Fuzzy C-Means clusters the group of customer who is having largest membership value in a particular group accordingly.

So in our experimental analysis, we clustered the customers on the basis of service used by a particular customer month wise. Hence we can get the clear view of churned services by the customers. At a time we can do the analysis of only two services of customers. Because of which it’s possible for us to represent the clusters graphically. Hence the user of our system can better understand the customer behavior and can use it to improve company performance. Also finally can easily find the service churn ratio.

Hence our FCM algorithm i.e. soft clustering has given accurate ratio of churned services with the use of membership value.

V. CONCLUSION

Hence we have observed that FCM has given a great support to the Customer Relationship Management. Hence this FCM algorithm can be used further to work on real time data sets of large size.

REFERENCES