CustomerRelationship and Management System

Aqueel Jivan[1], Rafsan Shaikh[2], Ibrahim Broachwala[3], Pritam Mandlik[4], Prof Nafisa Mapari[5]

[1,2,3,4] Final Year Students of Department of Computer Engineering, MH Saboo Siddik COE, MUMBAI-400008, INDIA.
[5]Assistant Professor, Department of Computer Engineering, MH Saboo Siddik COE, MUMBAI-400008, INDIA

Abstract: In the modern business landscape, it has become essential for organizations to monitor and address the requirements of their customers. A Customer Relationship Management (CRM) is a system which aims at increasing and retaining existing customers by creating business strategies that are based on the analysis of the customers' historical data. This paper discusses the data mining methods that are used for analysis such as classification, clustering and regression and introduces the concept of clustering using fuzzy logic. It describes fuzzy c-means approach which uses the membership value of data objects to cluster them and how it overcomes the drawbacks of traditional clustering methods.

I.INTRODUCTION

In today's competitive market, it is crucial to have a customer-oriented approach to business to ensure that the needs and requirements of customers are met. A Customer Relationship Management (CRM) system aims to provide an efficient analysis of customers data and an effective approach to develop their business strategies and processes which meet customer requirements. CRM refers to strategies, practices and tool which is used to manage and analyze customer's data and interactions, with the aim of improving meet customer expectation. CRM refers to strategies, practices and tool which is used to manage and analyze customer's data and interactions, with the aim of improving Relationship with customers [1]. Data which is collected from customer transactions becomes a very important source of information for organizations. Analysis is made to determine the product choice of customers and use the knowledge to increase the business. CRM is focusing directly towards acquisition of new customers, retaining the old customers and increase the sales growth. The extraction of hidden patterns of data with the help of different data mining methods can be classified into two types: Description methods and prediction methods. The data description methods focus on understanding and interpreting the data with the help of examples and the way in which the underlying data relates to its parts. The aim of the prediction-oriented models is to construct a behavioral model with new samples which can predict the values which are related to the sample [2].

II. EXISTING METHODS

1. DATA MINING

The process of Data Mining is of searching for patterns of interesting and hidden (hidden pattern) of a large data set stored in a database, data warehouse, or other data storage [3]. Data mining, often known as knowledge discovery in database (KDD) there are three types used to group objects into identified classes such as classification, regression and cluster.

1.1 CLASSIFICATION

Classification is used to allocate the information into any one of the predefined classes. Classification consists of two phases: The first is the training phase in which a portion of the dataset is used for training the classification algorithm which uses the attributes called features to determine the classes or labels. The second is the test phase which determines the accuracy performance of the classification patterns. Classes can be either binary or multi-valued.

1.2 CLUSTERING

Clustering is an important technique through which object grouping can be done (like the different groups of customers). The objects belonging to the same cluster are similar but those which are in the different groups are different. In this descriptive task a finite set of clusters are determined which identify or describe the data.

1.3 REGRESSION

Assuming a linear or nonlinear model of dependency, regression analysis can be used by us to predict the value of given (continuous) features based on the other features in the data [6]. Regression analysis is a statistical measure used to show the relationship between one or more free factors and dependent factors.

2. FUZZY CLUSTERING

2.1 FUZZY LOGIC

Unlike the conservative Boolean logic theory of computing in which the value of a base maps to only one of the values between 0 and 1, which means it can be either true or false. On the other hand, Fuzzy Logic is a form of multi-valued logic which takes into consideration the degrees of truth which range from 0 which means false to 1 which means true. The truth value of a base is determined as per the scale illustrated below:

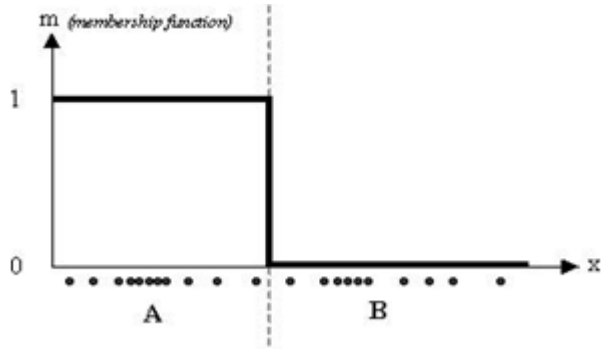
Table 1 Typical scale for Fuzzy truth values Fuzzy truth value Semantic meaning 0 false

0	False
0.1	Nearly false
0.2	very false
0.3	somewhat false
0.4	more false than true
0.5	as true as false
0.6	more true than false
0.7	somewhat true
0.8	very true
0.9	nearly true
1	True

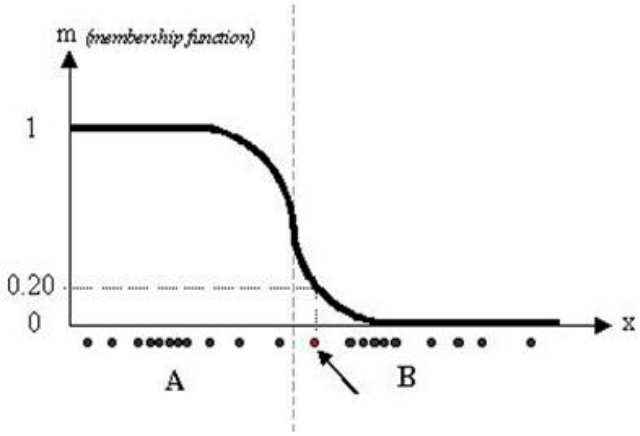
3. Categorization using fuzzy logic

Customer segmentation is the separation of all potential and existing customers based on a number of relevant attributes called features in internal homogeneous and other heterogeneous subsegments of customer segmentation [5]. Since customer information is not always sharp, the direct correlation to one specific segment is not practical. There are therefore several approaches that have been developed which use the concept of Fuzzy logic for the purpose of customer segmentation [7]. Clustering is the data mining procedure of grouping a set of data items into various clusters or groups so that objects inside the cluster have high similarity and are extremely dissimilar in alternate groups [8]. The advantage of using Fuzzy logic in the context of clustering is the development of intermediate classifications. A data-object belongs to any given cluster with a specific degree of membership. This has an advantage over conventional clustering methods like k-means, in which the algorithm generates binary or "crisp" classifications. Thus, a data-object or in the case of a CRM, a customer can only be allocated to a single cluster at a time. This could potentially lead to incorrect assumptions, because a customer can have characteristics that allow him to be in multiple clusters [7]. To understand better, we can consider this simple mono dimensional example. Let us represent a dataset scattered on an axis as shown in the figure below:

In the above fig, we can identify two clusters in closeness of the two data concentrations. Let us denote them using the notations 'A' and 'B' respectively. In conventional approach to this problem using the k-means algorithm, we associate each datum to a explicit centroid; thus, this membership function looked like the following:



In the approach using fuzzy logic called the Fuzzy C means clustering (FCM), instead the same given datum does not belong entirely to a well-defined cluster, but it can be positioned middle way. The membership function as a result tracks a smoother line which implies that a datum may belong to multiple clusters with varying values of the membership coefficients.



III PROPOSED METHOD

Fuzzy C means

Fuzzy C-Means is a data clustering technique based on fuzzy membership functions in which the existence of a data-object in a given cluster is determined by its degree of membership [9]. It is created on minimization of the subsequent objective function:

$$J_m = \sum_{i=1}^{n}\sum_{j=1}^{c} u_{ij}^{m} \left\| x_i - c_j \right\|^2 \qquad 1 \le m < \infty$$

where m is any real number greater than 1, u is the degree of membership of x in the cluster j, x_i is the ith of d-dimensional measured data, c is the d-dimension middle of the cluster, and ||*|| is any rule expressing the likeness among any measured data and the center.

The objective function above is optimized iteratively and it supports the fuzzy partitioning, with the update of membership u_{ij} and the cluster centers c_j by:

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}}$$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

This iteration will stop when $\max_{ij}\left\{\left|u_{ij}^{(k+1)} - u_{ij}^{(k)}\right|\right\} < \varepsilon$, where ε is a end condition between 0 and 1, whereas *k* is the iteration number. This technique meets to a local minimum or a saddle point of *J*.

Following are the steps of the algorithm.

1. *Initialize* $U=[u_{ij}]$ *matrix,* $U^{(0)}$
2. *At k-step: calculate the centers vectors* $C^{(k)}=[c_j]$ *with* $U^{(k)}$

$$c_j = \frac{\sum_{i=1}^{N} u_{ij}^{m} \cdot x_i}{\sum_{i=1}^{N} u_{ij}^{m}}$$

3. *Update* $U^{(k)}$, $U^{(k+1)}$

$$u_{ij} = \frac{1}{\sum_{k=1}^{C}\left(\frac{\|x_i - c_j\|}{\|x_i - c_k\|}\right)^{\frac{2}{m-1}}}$$

4. If $\| U^{(k+1)} - U^{(k)} \| < \varepsilon$ *then STOP; otherwise return to step 2*

4. CONCLUSION

This paper revisits the idea of a Customer Relationship Management systems which uses user data to obtain insights about the customers transactional behavior which is in turn used to improve an organization's relationship with the user. It describes the methods used for the same which include classification, clustering and regression. This paper proposes the use of a clustering method which makes use of fuzzy membership functions for the purpose of clustering data objects. This method has an advantage over traditional clustering methods like k-means where the segmentation of data is crisp and binary. In the fuzzy C-means approach a single data item can belong to multiple clusters based on its membership value.

5.REFERENCES

[1] R.Siva Subramanian, Dr.D.Prabha "A survey on customer relationship management",

[2] Ruxandra PET RE, "Data Mining Solutions for the Business Environment", Database Systems Journal vol. IV, no. 4/20 I 3

[3] P. A. Tan, M. Steinbach, and V. Kumar, Introduction to Data Mining .2006.

[4] Espin, R., Fernandez, J., Marx-Gomez, J., Lecich, M.: Compensatory Fuzzy Logic: A Fuzzy normative model for decision making. Investigación Operativa. Universidad de la Habana (2), 188–197 (2006)

[5]Freter, H.: Marktsegmentierung. Kohlhammer Verlag, Stuttgart (1983)

[6] Mishra, B. K., 1, Hazra, D., 2, Tarannum, K., 3, & Kumar, M., 4. (n.d.). 2016 International Conference System Modeling & Advancement in Research Trends (SMART). Moradabad: IEEE.

[7] Sven Kölpin and Daniel Stamer,, Categorization of Unlabeled Customer-Related Data Using Methods from Compensatory Fuzzy Logic, Soft Computing for Business Intelligence

[8] Sivaramakrishnan R Guruvayur, Dr.Suchithra R, a detailed study on machine Learning techniques for data mining, International Conference on Trends in Electronics and Informatics.

[9] Lisna Zahrotun, Implementation of Data Mining Technique for Customer Relationship Management (CRM) on Online Shop Tokodiapers.com With Fuzzy C-Means Clustering, 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE).