Customer Relationship Management Using Clustering And Classification Technique

Apurva Sharma[1], Dr. Harmaninder Jit Singh Sidhu[2]

[1]Research Scholar, Desh Bhagat University, Punjab
[2]Assistant Professor, Desh Bhagat University, Punjab

Abstract: *CRM (Customer Relationship Management) is a method and tool that helps organizations to maintain customer relationships in a structured manner. It can help to choose the right people or decide on new products that their customers might be interested in. Customer satisfaction plays an important role in any organization to maintain CRM. The main aim of this research is to improve CRM of an organization with cluster analysis. Similarity index has been used to determine the relationship between customers. On the basis of similarity index values, K-mean clustering approach has been used to form groups of various customers on the basis of their interest. After, clustering approach, classification technique such as SVM (Support vector machine) is used to provide the data as per customers need. To determine the efficiency of the proposed model, different parameters such as precision, recall, and F-measure have been measured.*

Keywords: *CRM, Similarity Index, Jaccard similarity, Precision, Recall, and F-measure*

I. Introduction

CRM (Customer Relation Management) is the business philosophy for acquiring and maintaining customers, increasing customer value, fidelity and preservation, and implementing customer-oriented strategies [1]. CRM, designed to enhance customer relationships, focuses on a comprehensive picture of how customer integrity, customer expectations, expectations, and behaviors are analyzed from a customer's event. Businesses can shorten the sales cycle and increase customer loyalty to improve a closer customer relationship and increase more revenue with good CRM [2]. This way, an excellent CRM can help businesses keep existing customers and attract new ones. Companies apply some of the ways to streamline customer relationships, including customer relationship management, customer value analysis, business strategy, and positive service mechanisms. In addition, companies strengthen the marketing and sales effectiveness of building good CRM. It is estimated that it will take five times as much on attracting new customers as it is to maintain the current company, according to the American Executive Board survey, and this relationship is particularly apparent in the service sector [3]. That's why companies understand how important it is to develop a good close relationship with existing and new customers. Instead of attracting new customers, they want to enable more business to their customers to keep existing customers and build a long-term customer relationship. For this reason, this study ensures that companies need to implement customer analysis to understand their customers, retain valuable customers, and eventually bring a lot of profits to themselves [4].

II. Material and Methods

Similarity index in CRM

The similarity is the amount that reflects the strength of the relationship between two customer relations, which represents the degree of similarity between the two customers [5]. In this research work, to find the similarity between customers, Jaccard similarity has been used. The Jaccard similarity index (sometimes referred to as the Jaccard similarity coefficient) compares two groups of members to see which members are shared and which members are different. Jaccard similarity can be calculated as:

$$\text{Jaccard index} = \frac{\text{data in both sets}}{\text{data in either set}} * 100 \quad (1)$$

On the basis of similarity, clustering process has been carried out. In clustering, the data with similar value form the group with similar data objects together [5]. This measure of similarity is the most common and is based on distance functions in most applications to group objects in the cluster. Clustering is formed in such a way that any two data objects within a cluster have a minimum distance value, and any two data objects across different clusters have a maximum distance value [6]. Clustering using distance functions (called distance-based clustering) is a very popular technique for clustering objects and has given good results.

CRM with clustering

Cluster mining is widely used for CRM. It provides business listings that enable businesses to offer their customers special, personalized services and products [7]. In a commercial environment, clustering is used, for example, in the following areas:

1. Cross-marketing
2. Cross-selling
3. Customization of marketing plans for varied customer types
4. To decide what media method to utilize
5. Identify the shopping goals

Clustering analysis is a data mining technique that delivers data objects to anonymous groups of high-level objects [8]. Clustering is the task of segmenting a diverse population in a number of more homogeneous clusters. Clustering algorithms are classified into partition or hierarchical.

K-means clustering

K-means clustering is used in cluster observations in groups of relevant observations without any prior knowledge of those relationships [9]. In this paper, we have used k-means algorithm to customer segments based on transaction records. This algorithm aims to assign a set of n data objects to clusters in k to achieve a high uniformity of intra-cluster and a low resemblance of the same cluster.

It set observations to k groups, where k is provided as an input parameter. It then assigns each observation to clusters based on the proximity of observation to the cluster definition [10]. Cluster mean means to be recomputed and the process starts again. The algorithm selects k points as initial cluster centers. Finally, this algorithm aims to minimize an objective function, in this case, a single error function. The objective function is defined below:

$$J = \sum_{j=1}^{k} \sum_{i=1}^{n} \left\| x_i^j - c_j \right\|^2 \quad (2)$$

As shown in the above equation, $\left\| x_i^j - c_j \right\|^2$ is the selected distance metric between the data point and the cluster center, an indicator of the distance of the n data points from the respective cluster centers [11]. The algorithm consists of the below steps:

Step 1: Put the K point into the space shown by the cluster object. These are expressed as the initial group centroid.

Step 2: Assign every object to the group with the closest centroid.

Step 3: After all the objects have been assigned, re-compute the position of the K centroid.

Step 4: Repeat steps 2 and 3 till the centroids no longer move. This develops a group that collects objects into groups from which the parameters to be reduced can be computed.

Classification process

Based on the clustering process, customer relationship has been classified using SVM (Support vector machine) approach [12]. SVM is a binary classification approach which is used to differentiate between two types of data. The data is separated by using hyper-plane.
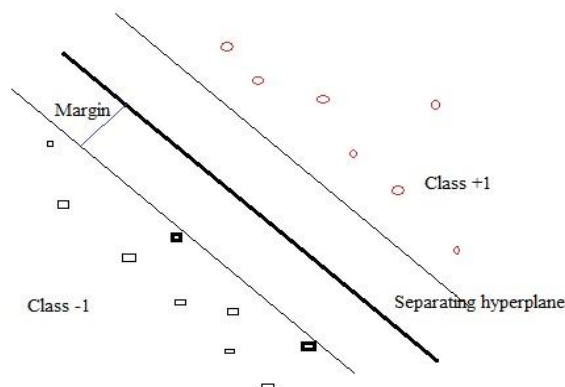

Fig no 1: Support vector machine working

As shown in the above figure, the hyperplanes are separated by two types of data. One data type is represented as a circle in red and another by the square in black color [13]. A margin has been set for both data types.

I. Related work

Although CRM is debatable in studies focused on various industries over the past two decades (Croteau & Li, 2003; Sigala, 2005), many disagreements remain relevant to its meaning. Some of these studies are related to the retrieval and long-term retention of customers (Ling & Yen, 2001), while others consider it a business strategy that provides continuous integration of each customer-centric business area (Sathish, Shan, Raman, 2002). According to Payne (2006), CRM is a business and marketing strategy that integrates technology, processes, and business activities around the customer. Some additional studies have investigated CRM in the opportunities of providing it for maximizing revenue, for example by attracting and maintaining "economically important" customers while removing "incredible value" (Pan & Lee, 2003). Bradshaw and Brash (2001) have considered CRM to be a management method that enables organizations to recognize, attract, and supplement the retention of useful customers by managing relationships, contact them. Marketing researchers, management, and information systems have made various attempts to identify CRM from three different views of philosophy, strategy and technological solutions. Ryals and Knox (2001, p.535); Zablah et al. (2004, pp. 478) has defined CRM as the framework of business philosophy. Rababah, Mohd, and Ibrahim (2010, pp. 223) have analyzed the best definition of CRM that would include all these views within a comprehensive view. Accordingly, the authors have referred to CRM as "the establishment of a customer-focused culture in which a strategy was created for the acquisition, enhancement of the profitability, and retention of customers, enabled by an IT application; achieving both benefits for both organizations and customers".

II. Proposed Methodology

In this section of the paper, we have described the implementation methodology of the proposed work with the concept of SVM as a classifier. The implementation methodology of the proposed work is given as;

Step 1: Design a framework for the simulation of proposed work to implement a structured CRM system.

Step 2: Develop a code for documents uploading in proposed work for further simulation. Step 3: Apply pre-processing of uploaded documents to find out the feature sets using the concept of data normalization. The used pre-processing algorithm is given as;

Pre-processing Algorithm
Input: Document files (D)
Output: the Weighted value of documents
Calculate the size of documents → [Row, Columns]
For I = 1 → Row
For J = 1 → Columns
  N_documents = Normalization (D (I, J))
  W_documents = Weight (N_documents (I, J))
 End
End
Return; Weighted value of documents = W_documents
End

After the pre-processing of uploaded documents, Jaccard similarity is applied on weighted value by using step 4
Step 4: Design a Jaccard similarity algorithm to find out the similarity between uploaded documents based on their weighted value. The Jaccard similarity algorithm is given as:

Jaccard Similarity Algorithm
Input: Weighted data
Output: Similarity Index of documents
Calculate Intersection of data=x∩y
Calculate Union of data=x∪y
Jac_sim = [];
Sim_count = 0;
For m = 0 → data.length
  Current_ data = data (m);
  For n = I +1 → docs.length
    L= |X∩Y| / |X∪Y|
    Jac_sim[sim_count,0] = current_ data;0
    Jac_sim[sim_count,1]= docs(n);
    Jac_sim[sim_count,2]=L;
    Sim_count = Sim_count + 1;
  End

End
Return; Jac_sim as a Similarity Index
End
In this algorithm, data is the weighted input whereas; Jac_sim is the output of algorithm and on the basis of Jac_sim index we apply k-means clustering algorithm.
Step 5: After that, develop a code for K-means clustering algorithm and find out the clusters of uploaded documents. The k-means clustering algorithm is given as;

K-means Clustering Algorithm
Input: Similarity Index (SI)
Output: Clustered Data
Define centroids of clusters, C1, C2 …Cn
Calculate the size of Similarity Index → [Row, Columns]
For I = 1 → Row
For J = 1 → Columns
  Cluster 1= Distance (C1, SI (I, J))
  Cluster 2= Distance (C2, SI (I, J))
       .
       .
       .
  Cluster n= Distance (Cn, SI (I, J))
    End
End
Return; Clustered Data = Cluster 1, Cluster 2…….Cluster n.
End
After the clustering, we apply SVM as a classifier to design a structure CRM system.
Step 6: Apply SVM for the training of the proposed CRM system using the given algorithm
SVM Algorithm
Input: Weighted value as a Training Data
Output: Classified Results
Initialize parameters   – Kernel function (Gaussian or RBF)   – Training Data (T)
Based on Training Data define Groups
For 1 → all Training Data
   If Training Data belongs to Customer relationship
     Group 1 = Customer satisfied
   Else
     Group 2 = Customer not satisfied
   End
End
SVMStruct = SVMTRAIN (Training Data, Group, Kernel function)
Test data = Weight of Current Data
Return; Classified output = SVMCLASSIFY (SVMStruct, Test data)
End
On the basis of the above-mentioned methodology, the work has been simulated and the following results have been obtained.

III. Result and Discussion
This section has explained the results obtained after the simulation of the proposed work. For the simulation, parameters such as Precision, Recall, and F-measure has been measured.

Table no 1 Performance parameters

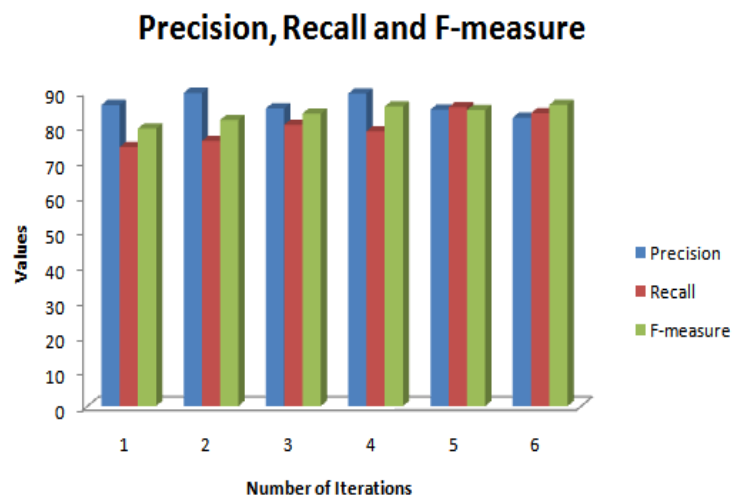| Iterations	Precision	Recall	F-measure
1	86.23	74.18	79.49
2	89.69	75.90	81.95
3	85.26	80.55	83.77
4	89.50	78.71	85.80
5	84.90	85.62	84.79
6	82.60	83.89	86.29

Precision, Recall and F-measure

Fig no 2: Computation of Precision, Recall, and F-measure

Above figure and table represents the computation for precision, recall, and F-measure. As depicted, the X-axis depicts the number of iterations and Y-axis defines the values obtained after the simulation. Red bar in the graph defines the results of recall, blue bar defines the result of precision and green bar defines the results of F-measure. SVM has been used as a classifier for customer relationship management. The average value of Precision is 86.36, Recall is 79.80 and F-measure is 83.68.

IV. Conclusion

CRM is a system designed to improve relationships with existing customers, find new prospects and win back to previous customers. The system can be implemented in software that facilitates the collection, organization, and management of customer information. The CRM strategy is beneficial to both small and large-scale commercial enterprises. One of the main benefits of using CRM is to get better customer satisfaction. By using this strategy, all transactions involving the provision of services, marketing and sales of products to customers can be performed in an organized and systematic manner. Following a CRM strategy helps build better communication within the company. This research has utilized the concept of clustering to improve CRM. For determining the relation between the customers, the similarity index has been used. For the development of group as per customer's interest, K-mean clustering approach is used and after that, SVM is used as a classification technique to provide the data as per customer requirement. Parameters, such as Precision, Recall, and F-measure have been used to determine the effectiveness of the proposed work. The average value of Precision is 86.36, Recall is 79.80 and F-measure is 83.68.

References

[1]. Croteau, A.-M., & Li, P. (2003). Critical success factors of CRM technological initiatives. Canadian Journal of Administrative Sciences, 1(20), 21–34.
[2]. Sigala, M. (2005). Integrating customer relationship management in hotel operations: Managerial and operational implications. International Journal of Hospitality Management, 24(3), 391–413.
[3]. Ling, R., & Yen, D. C. (2001). Customer relationship management: An analysis framework and implementation strategies. Journal of Computer Information Systems, 41(3), 17.
[4]. Sathish, S., Pan, S., & Raman, K. (2002, August 9–11). Customer relationship management (CRM) network: A new approach to studying CRM. AMCIS 2002 Proceedings, paper 79, Dallas, TX.
[5]. Payne, A. (2006). Handbook of CRM: achieving excellence in customer management. Oxford: Routledge
[6]. Pan, S. L., & Lee, J.-N. (2003)). Using e-CRM for a unified view of the customer. association for computing machinery.
[7]. Bradshaw, D., & Brash, C. (2001). Managing customer relationships in the e-business world: How to personalize computer relationships for increased profitability. International Journal of Retail & Distribution Management, 29(12), 520–530
[8]. Ryals, L., & Knox, S. (2001). Cross-functional issues in the implementation of relationship marketing through customer relationship management. European Management Journal, 19(5), 534–542.
[9]. Zablah, A. R., Bellenger, D. N., & Johnston, W. J. (2004). An evaluation of divergent perspectives on customer relationship management: Towards a common understanding of an emerging phenomenon. Industrial Marketing Management, 33, 475–489
[10]. Rababah, K., Mohd, H., & Ibrahim, H. (2010). A unified definition of CRM towards the successful adoption and implementation. In the 3rd Lifelong Learning International Conference (3LInC'10), Kuala Lumpur, Malaysia.
[11]. Kaul, D. (2017). Customer Relationship Management (CRM), Customer Satisfaction and Customer Lifetime Value in Retail. Review of Professional Management, 15(2), 55-60.
[12]. Eichorn, F. L. (2018). Internal customer relationship management (IntCRM) a framework for achieving customer relationship management from the inside out. Management, 2, 1.
[13]. Rahimi, R., & Kozak, M. (2017). Impact of customer relationship management on customer satisfaction: The case of a budget hotel chain. Journal of Travel & Tourism Marketing, 34(1), 40-51.

[14]. Deb, S. K., Jain, R., & Deb, V. (2018, January). Artificial Intelligence—Creating Automated Insights for Customer Relationship Management. In 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence) (pp. 758-764). IEEE.

[15]. Ascarza, E., Ebbes, P., Netzer, O., & Danielson, M. (2017). Beyond the target customer: Social effects of customer relationship management campaigns. Journal of Marketing Research, 54(3), 347-363.

[16]. Diffley, S., McCole, P., & Carvajal-Trujillo, E. (2018). Examining social customer relationship management among Irish hotels. International Journal of Contemporary Hospitality Management, 30(2), 1072-1091.

[17]. Badwan, J. J., Al Shobaki, M. J., Naser, S. S. A., & Amuna, Y. M. A. (2017). Adopting technology for customer relationship management in higher educational institutions. International Journal of Engineering and Information Systems (IJEAIS), 1(1), 20-28.

[18]. Rahimi, R. (2017). Customer relationship management (people, process and technology) and organisational culture in hotels: Which traits matter? International Journal of Contemporary Hospitality Management, 29(5), 1380-1402.

[19]. Rahimi, R., Köseoglu, M. A., Ersoy, A. B., & Okumus, F. (2017). Customer relationship management research in tourism and hospitality: a state-of-the-art. Tourism Review, 72(2), 209-220.

[20]. Lemma, M., & Kebede, G. (2018). 6.2 Effects of Customer Relations Management on Market Performance in the Banking Sector: A Comparative Study of Selected Private Banks in Ethiopia.

[21]. Rahimi, R. (2017). Organizational culture and customer relationship management: a simple linear regression analysis. Journal of Hospitality Marketing & Management, 26(4), 443-449.

[22]. Amuna, Y. M. A., Al Shobaki, M. J., Naser, S. S. A., & Badwan, J. J. (2017). Understanding Critical Variables for Customer Relationship Management in Higher Education Institution from Employees Perspective.