Customer Relationship Management Tool

*Ashtami Bhuleskar, †Vibhati Joshi, ‡Sanika Save
*Fr. Conceicao Rodrigues College of Engineering, Bandra (W), Mumbai-50
†bhuleskar.ashtami@gmail.com, vibhatij@gmail.com, sanikapsave@gmail.com

***Abstract*—Abstract A huge amount of research has been done in the field of Customer Relationship Management in the recent years from a business point of view. These efforts have been towards determining factors of CRM, efficiency parameters, how it affects market trends and company profits, etcetera. The fields of technology and business in CRM have been fairly disparate in most of these efforts. This project aims to bridge this gap of the technical implementation and its subsequent business implications by integrating them together in a fulledged CRM solution. It is an attempt to take the middle path between software development and management issues and delicately balancing them both in a unique two way system that fulfils the criteria of both the company and the customer.This software aims to perform analysis of market trends and subsequent customer habits also finding new potential profitable customers for the company using data mining . On the whole, the project offers a proficient and dynamic merger of the fields of technology and business perspectives. Thus, it provides a wholesome solution to tackle customer relationship strategies thereby making the model as profitable for the company and the same time gaining high levels of customer satisfaction.**

***Index Terms*—Customer Relationship Management, Clusters, Extraction, Transaction Database, Association Rules, Support Count, Data Mining, K-means, Apriori algorithm.**

I. Introduction

Data mining also known as Knowledge Discovery in Database(KDD). The purpose of data mining is to abstract interesting knowledge from the large database. From the analysis of abstracted patterns, decision-making process can be done very easily. [1] This paper presents a technique and an automated tool for Customer Relationship Management which is used to make more efficient business-customer relationships in order to maximize client satisfaction and thereby improve customer loyalty and retention. Data mining techniques are useful to analyze the customer behavior. Project is designed to find existing customers which can be more profitable for the company if tracked in a proper way. [2] This can be achieved by analyzing their product usage habits. Such frequently appearing combinations can be termed as common behaviour which can be used to track down other customers. Also the other objective is to target people of the region under consideration based on the existent schemes of bank. These people can be future aspirants of the bank. The project focuses on combining the business aesthetics of the customer relationship module along with its actual technical implementation. It aims at using simple methodologies and data mining algorithms to effectively extract and analyze data. [3] [4]

II. CUSTOMER RELATIONSHIP MANAGEMENT

Data mining creates the ideal environment for making customer relationship management (CRM). The approaches of how data analytics can be used to make various CRM functions like customer segmentation, communication targeting, retention, and loyalty much more effective are present and implemented before. [5] CRM is considered as the database marketing of an organization with the database of customers. CRM identifies the most profitable customers and provides the highest level of service to them. Customercentric approach helps to improve customer satisfaction and thereby maximize customer retention. Good customer relationship management software helps you to understand customer expectations and provides services. [14]

III. PROPOSED WORK

We are planning to combine the intricacies of the Information Technology along with the subtleties of the business perspectives. Our project aims at comparing how a business model that inculcates CRM fares better in terms of profits, market trends and customer satisfaction than a traditional supply driven model that ignores the customer. [6]

A. Detail extraction

The user is provided with an easy accessibility to customer information based only on the account number as the input. For this, the UI has been made that provides a text box to accept the input from the user. This input has to be a valid account number. Once the account number has been entered, clicking on the get details button will retrieve the corresponding customers personal details as well as the schemes that he is using within the bank domain. This enables the manager or end user to get a comprehensive knowledge about each person and how valuable he can be to the bank. [7]

B. Customer Tracking

1) Generation of Transaction Database

In this paper, the fundamentals of the Apriori algorithm in order to find frequently occurring item-sets and strong associations of existing customers product patterns. [8] [9] In order to apply an algorithm of this sort, we need to first have a transaction database ready. Normally Apriori algorithms take a transaction database as the input in order to create associations.Considering a real world scenario of a banking system, it is not conventional to have a transaction database present in the system.Also, Apriori results and the associations generated cannot be predicted. This was a major setback in terms of getting haphazard results that might be of little or no use to the user as it would prove to be a trial and error

mechanism.

To overcome this issue, user gets select at the most three schemes that he necessarily wants to be a part of his associations that are formed. This helps in streamlining the results according to the needs of the user. Selecting these fixed tables helps the user to determine which schemes he is analysing and thereby provides him with a clear view of what the application has achieved for him. This completely removes the uncertainty and trial and error nature of the Apriori algorithm. These fixed tables now greatly influence the creation of the transaction database on which the final algorithm must be implemented to generate the associations of customer trends.

2) Implementation of Apriori on the generated transaction database

Transaction database is taken as input for the algorithm. The transaction database basically represents in each tuple an account number and the subsequent schemes that the customer has enrolled for with the bank. This then serves as the base for the further implementation of the algorithm. Keeping the bank system in mind, we have designed the application to give associations of maximum five schemes combination. [10] Support count value is then taken as a standard against which all the combinations formed will be compared in order to determine whether they pass or not. Thus the application terminates in the case of following conditions:

- If no more combinations can be formed for new associations.
- If no combinations cross the specified support count.
- If the loop finishes iterating for 5 number of times.

In the banking system, we had to keep in mind that the failed schemes at each stage were also as important as the ones that pass. Thus we propose to store both passed and failed schemes. This assists the user in a dual way by helping him boost weak scheme associations and gain profits by further enhancing existing strong associations. [11]

Apriori algorithm represents the candidate generation approach. It generates candidate (k+1) itemsets based on frequent k-itemsets. Apriori is a Breadth First Search Algorithm (BFS).

C. Lead Capture

Clustering analysis is a data mining technique that maps data objects into unknown groups of objects with high similarity. Clustering is the task of segmenting a heterogeneous population into a number of more homogenous clusters . Clustering algorithms are classified into partitional or hierarchical. [12]

The k-means clustering is used to cluster observations into groups of related observations without any prior knowledge of those relationships. In this paper, we apply the k means algorithm to segment potential customers based on their age. This algorithm aims to assign a set of n data objects to k. [13]

Definition : The support (s) for an association rule X ∈ Y is the percentage (%) of transactions in the database that contains X U Y. support (X ∈ Y) = P (X U Y), where P is the probability.
Input: D: Database of transactions;
minsup: minimum support threshold
Output: L: frequent itemsets in D
Method:

```
1:  for each itemset l1∈Lk-1 { do
2:    for each itemset l2∈ Lk-1 { do
3:      if (l1 [1]= l2 [1]) AND (l1 [2]= l2 [2]) AND (l1
        [k-2]= l2 [k-2]) AND (l1 [k-1]≤ l2 [k-1]){ then
4:        c=l1 infl2;
5:        for each itemset l1∈k-1 { do
6:          for each candidate c ∈ k { do
7:            if l1 is the subset of c then
8:              c.num++ }}}}};
9:            end if
10:         end for
11:       end for
12:     end if
13:   end for
14: end for
15: C'k=( c ∈ Ck OR c.num=k) ;
16: return C'k;
```

Algorithm 1: Procedure apriori_gen(Lk-1:frequent(k-1)-itemsets)

```
1:  L1=find_frequent_1-itemsets(D);
2:  for (k=2;Lk-1 ≠ ϕ; k++){ do
3:    Ck=apriorigen(Lk-1, min_sup);
4:    for each transaction t ∈ D do
5:      Ct=subset(Ck,t);
6:      for each candidate c ∈ Ct
7:      c.count++;
8:      }
9:    end for
10:   Lk={ c ∈ Ck | c.count ≥ min_sup };
11:   if (k ≥ 2){ then
12:     delete_datavalue(D, Lk, Lk-1)
13:     delete_datavalue(D, Lk, Lk-1)};
14:   end if
15:   }
16: end for
17: return L=UkLk ;
```

Algorithm 2: Apriori Algorithm

clusters in order to achieve a high intracluster similarity and a low inter cluster similarity.

The algorithm clusters observations into k groups, where k is provided as an input parameter. It then assigns each observation to clusters based upon the observations proximity to the mean of the cluster. The clusters mean is then recomputed and the process begins again. The algorithm arbitrarily selects k points as the initial cluster centers.

The algorithm is composed of following steps:

k: the number of clusters
Input: D: Dataset containing n objects;
Output: A set of K clusters.
Method:

1: *arbitrarily choose k objects from D as the initial cluster centers* ;
2: *repeat*;
3: *(re)assign each object to the cluster to which the object is the most similar,based on the mean value of the objects in the cluster*;
4: *update the cluster means, that is, calculate the mean value of the object for each cluster*;
5: *until no change*;

Algorithm 3: Algorithm For K-Means

IV. CONCLUSION

Analysis of customer behavior enables companies to improve support of their customer oriented business processes, which aims to improve the overall performance of the enterprise. This paper also focuses on getting more customer satisfaction.

Data mining methodology has a tremendous contribution for researchers to extract the hidden knowledge and information The research described in this paper also identified significant product association rules. Product association rules can be used to motivate customers to increase their scheme acqusition and keep loyal to the company. The behavior of customers can be easily identified.

Apriori algorithm is very helpful to find the best associations. The most frequent itemsets can be easily found out from the database.

K-means algorithm helped to generate clusters of the data points available in the databse arounf the centroids considering age as a parameter.

ACKNOWLEDGMENT

We would like to thank the project guide and other faculty members for their continuous support and guidance also for their detailed comments on earlier versions of this paper.

REFERENCES

[1] L. S.-g. L. J. C. Han-joon, Kim TaeHee, "Automated data warehousing for rule-based crm systems," Adelaide, Australia, 2003.
[2] D. P. G. Yongping, "A new improvement of apriori algorithm for mining association rules," in *Computer Application and System Modeling (ICCASM), 2010 International Conference*, vol. 2, October 2010, pp. 529–532.
[3] L. X. E.W.T. Ngai and D. Chau, "Application of data mining techniques in customer relationship management: A literature review and classification, expert systems with applications," Ph.D. dissertation, 2009.
[4] J. T. Dawn E. Holmes and L. C. Jain, "Data mining techniques in clustering, association and classification, data mining: Foundations and intelligent paradigms, intelligent systems reference librar and trusted computing (iptc)," vol. 23, Springer, 2012, pp. 1–6.
[5] F. Buttle, "Customer relationship management: Concepts and technologies," in *First edition, Oxford, Elsevier Butterworth Heinemann, ISBN*, 2009, p. 522.
[6] M. W. W. M. Sonja Grabner-Kraeuter, Gernot Moedritscher, "Performance monitoring of crm initiatives," in *Performance Monitoring of CRM Initiatives, in Proceedings: IEEE conference on System Sciences*, 2007.
[7] R. Agrawal, "Fast algorithms for mining association rules in large databases," *Proceedings of the ACM SIGMOD International Conference Management of Data*, pp. 207–216, 1993.
[8] Z. Y. Xu and C. Zhang, "An optimized apriori algorithm for mining association rules," *Comput. Eng., Shanghai*, vol. 29(19), pp. pp. 83–84, 2003.
[9] W. Y. X. W. E. W. B. Chen;, "The research of improved apriori algorithm for mining association rules," Nov.2008, pp. 513–516, 10–12.
[10] P. J. Han, J. and Y. Yin, "Mining frequent patterns without candidate generation," in *the 2000 ACM SIGMOD International Conference on Management of data*, Dallas, Texas , United States, 2000, pp. 1–12.
[11] W. J. Luo Ke, "Apriori algorithm based on the improved [j].computer engineering and application," in *proc. The Global Telecommunications Conference (GLOBECOM)*, vol. 20, 2001, pp. 20–22.
[12] M. R. G. Seyed Mohammad Seyed Hosseini, Anahita Maleki, "Cluster analysis using data mining approach to develop crm methodology to assess the customer loyalty, expert systems with applications," *10th International Conference*, vol. 37, pp. 5259–5264, 2010.
[13] O. Z. Ismail, R. and Bakar, "Associative prediction model and clustering for product forecast data, intelligent systems design and applications(isda)," 2010.
[14] E. Janny.C.Hoekstra, "The lifetime value concept in customer-based marketing, journal of market focused management [m]," vol. 3, 1999, pp. 257–274.