Application of Data Mining/Strategic CRM
for Indian Public Sector Banks

R. Udayakumar, K.P. Kaliyamurthie, Khanaa and A.V. Allin Geo

School of Computing Science, Bharath University, Chennai-73, India

Abstract: This Intensified competition and increased customer power have driven all financial institutions to pursue a variety of new strategies, all of which are dependent on new ways of understanding and relating to customers. Expansion without deep sophisticated customer intelligence can lead to unprofitable penetration of the wrong markets or the use of ineffective sales and service channels. As a result; customer set rather than product or geography increasingly defines market strategy. Elimination of product and geographic barriers enforces a customer focused strategy requiring data about customer lifestyle, life stage, net worth and personal interests dictate the portfolio of products, channels, level of service and technology infrastructure. The key to developing a multi-dimensional view is by understanding the overall relationship the organization has with the customer. This can be accomplished with the aid of a CRM solution that is easy to use and that accurately tracks all aspects of the relationship so that the customer receives a consistent experience no matter which interaction method he or she chooses. Rewards of CRM are significant in terms of productivity and profitability through Improved Customer loyalty and Satisfaction and reduction in Operating Costs. However, implementing a CRM strategy is challenging. Successful CRM projects require an effective mix of people, process and technology issues. This paper after identifying the customer attributes for CRM, for Indian Public Sector banks, analyzes through OLAP and Data Mining tools to develop effective strategies for step-by-step implementation of CRM to accurately track a holistic perspectives of the customer relationship.

Key words: Data Mining · CRM · Association Rules · Customer Segmentation · Customer Satisfaction Analysis · Next-Sequential-Product Analysis

INTRODUCTION

We now live in a society where the demands on business are so much greater than ever before. No business, unless it is a state monopoly, can stay in business without satisfied customers. Just as people cannot live without eating, corporations cannot continue to exist without satisfied customers. Moreover, due to the fact that customer expectations are constantly increasing, corporations are now required to go beyond their primary need of satisfying the customers, to exceed their expectations. Corporations therefore have to shift their customer focus from purely satisfying customers to create loyalty and trust through mutually beneficial, long-term relationships [1, 2].

Market competition demands corporations to continuously seek means to gain customer loyalty [3]. However, although corporations are realizing the value of keeping customers loyal, no one knows for sure how to do it. Corporations measure customer satisfaction and hope that if the satisfaction scores are good, the customers will stay with the firm. But the truth is that even satisfied customers leave for the temptation of competitors' offers.

Customer Loyalty and CRM: Loyal customers bring several advantages. They usually lead to increased revenues for the corporation, result in predictable sales and profit streams and are more likely to purchase additional goods and services [4, 5]. Furthermore,

Corresponding Author: R. Udayakumar, School of Computing Science, Bharath University, Chennai-73, India.

customers who are familiar with a brand are more likely to mention it to their friends and tend to be concerned in the feedback and evaluation of the product, which is critical in today's business environment. Loyal customers also tend to buy through alternative channels, for instance through the Internet, which might increase the total consumption and reduce the costs of doing business with them.

Customer loyalty moreover embraces many different concepts, for instance customer relationship management, customer retention marketing and one-to-one marketing [6]. These concepts are concerned with customer loyalty because of the benefits of retaining customers as well as the activities it involves, which aim at developing long-term relationships. Furthermore, most importantly for businesses, trust is a necessary condition for loyalty [7].

While retaining customer loyalty has been a sales principle for a very long time, Customer Relationship Management (CRM) is actually a tremendous step forward in creating a system that can provide a means for retaining individual loyalty in a world of nearly 6 billion souls. In order to understand CRM, you must also understand the changing nature of the customer because customers are not what they used to be.

CRM Strategies: According to Gefen and Riding (2002) CRM can be divided in to three different types: *operational, analytical and collaborative CRM.* The type of strategy is important as it indicates the tactics an organisation is adopting in implementing CRM [8].

Operational CRM, also called 'front-office' CRM, spans all the areas where direct customer contact occurs. These are customer 'touch points'. This class of CRM enables and streamlines communications to and from customers, but not necessarily optimizing the service [9].

Analytical CRM also called 'back-office' or 'strategic' CRM, involves comprehending the customer activities that had happened in the front office. Analytical CRM takes help of technologies (to compile and process the heaps of customer data to enable analysis) and new business processes to refine customer-facing practices to increase loyalty and profitability [10].

Collaborative CRM is a coordination network that provides the neural paths to the customer and supplier **(Schubert and Koch, 2002).** It could mean a portal, a partner relationship management application, or a customer interaction center. According to Fayerman, it

could also mean communication channels such as the Web or e-mail, voice applications, or snail mail. Fayerman further states that it also could mean channel strategies. According to Schubert and Koch, collaborative CRM is any CRM function that provides a point of interaction between the customer and the channel itself. Some important collaborative CRM are:

ECRM – enterprise CRM
PRM – Partner Relationship Management
cCRM – Collaborative CRM
SRM – Supplier Relationship Management

According to Greenberg (2001) the goal with CRM is to recognize and treat each customer as an individual using the three types of CRM.

Association Rule Mining for CRM: Data mining has attached a growing amount of attention in retail industry for improving strategies. Essentially the data mining tools tries to identify, extract and analyze potentially valuable data from numerous sources in a way that delivers useful information [11]. The primary data mining tasks include mining association rules, classification rules, summarization rules, clustering and similarity rules. Association rule mining finds interesting associations and/or correlation relationships among large set of data items. Association rules shows attribute value conditions that occur frequently together in a given dataset. A typical and widely-used example of association rule mining is Market Basket Analysis [12].

Methodology of Study: The methodology of study comprises of the following steps:

Identifying customer demagraphic, transactional etc. attributes Design questionnaires for data collection

Identify suitable customers in the selected branches of the banks Build customer aware about the benefits from the study

Issue to and receive filled questionnaires from the customers

Aggregate the data collected in the tables designed in MS Access / MS Excel Use Data mining tools to analyse customer behaviour patterns

Analyze the patterns to develop effective CRM

Develop a holistic frame work of CRM implementation for banks for customer loyalty and profitability.

Table 1: Customer Clusters Based on Age

Cluster No.	Min. Age	Max. Age	Density	Average Age
1	40	48	128	43.68
2	21	31	131	26.98
3	32	39	138	35.47
4	49	78	77	53.90

Data Analysis and Results: The OLAP analysis was performed on the data collected [13]. This was done to study the behaviour of the customers with respect to various attributes viz. age, income, channel utilisation, etc. The OLAP analysis provided an insight as to what the past behaviour of the customer was and thereafter along with data mining tools future behaviour of the customers may be predicted.

Customer Segmentation: Based upon the data collected from the customer feedback, the customers have been segmented based on their age and income group. The most suitable customer age brackets that are found using data mining clustering algorithm – PAM is shown in Table 1.

Channel Analysis: The various channels of banking services considered for the current analysis are:

- Physical Branch
- ATM
- Website/ Internet Banking
- Telebanking

Physical Branch: Main results of analysing Age vs. No. of Branch Visits are

5.	10.0	82.0	(Age)49-60=>(No_of_Br_visits)1
6.	21.0	71.0	(Age)40-48=>(No_of_Br_visits)1
7.	16.0	47.0	(Age)32-39=>(No_of_Br_visits)1
9.	10.0	48.0	(Age)21-31=>(No_of_Br_visits)0

From the results it is observed that customers within cluster 2 are the ones who visit the branches very less or almost never (Rule 9). The customers in cluster 3 and some customers in cluster 1 are the ones who like to visit the branches sometimes (Confidence 47% and 71% respectively). People in cluster 4 are the ones who are almost certain to visit the branches for transactions (Confidence 82%). These are usually the people on the older side, so it is understood that their liking for the face-to-face talk with the branch employees are more. As far as the cluster 4 is considered they would be difficult to be migrated from branch banking for they are the ones who love to come to the branch and have a chat with the employees while doing any transaction.

ATM: Main results of analysing Age vs. No. of ATM Visits are:

5.	12.0	40.0	(Age)49-60=>(No_of_ATM_visits)0
6.	16.0	53.0	(Age)40-48=>(No_of_ATM_visits)0
7.	16.0	47.0	(Age)32-39=>(No_of_ATM_visits)1
9.	14.0	66.0	(Age)21-31=>(No_of_ATM_visits)1

Customers within Cluster 3 are the ones who though show good response to ATM banking, yet since the confidence is only 47% they need to be encouraged more to use it. This result indicates that some of them who are exposed to the benefits of the ATM prefer using such channels while others are yet to migrate from branch banking. It would require some proper awareness programmes from the banks side to make these not yet IT savvy customers know about the benefits of using such channels. The customers in cluster 2 are the ones who can be safely assumed to use the ATM the most number of times. They command 66% confidence for such a statement. But the clusters 1 and 4 have the customers who are not at all interested in using the ATM channel for banking (Rule 5 and 6).

Internet Banking/Telebanking: Main results of analysing Age vs. No. of Internet Banking/ Telebanking are:

2.	74.0	91.0	(Web_visits)0=>(No_of_Telebank_visits)0
8.	10.0	89.0	(Age)49-60=>(No_of_Telebank_visits)0
9.	12.0	100.0	(Age)49-60=>(No_of_Website_visits)0
10.	26.0	87.0	(Age)40-48=>(No_of_Telebank_visits)0
11.	28.0	97.0	(Age)40-48 =>(No_of_Website_visits)0
12.	30.0	87.0	(Age)32-39=>(No_of_Website_visits)0

From the results it is evident that website visit and usage of telebanking facility are somewhat related directly (rule no 2). This may be because both are new age delivery channels and a customer not confident with one advance channel is not confident with the other advance channel also and vice-versa. It is also noticed that people in the clusters 1, 3 and 4 are not at all interested in these two channels. It is here that the banks have to go in for awareness programmes of these two delivery channels as these are low cost ones to the banks. People in cluster 2 are somewhat acquainted with the website banking, but not so much with telebanking. A reason for this may be due to the fact that many of the public sector banks do not have the telebanking facility at all.

Churn Analysis: The main results of churn analysis are:

19.	14.0	75.0	(Loans_Taken)3=>(Years_with_Bank)1
37.	13.0	68.0	(Benefit_of_this_Bank)1=>(Recommend)2
48.	14.0	81.0	(Loans_Taken) 4 => (Years_with_Bank) 2
81.	18.0	66.0	(Loans_Future) 1 => (Years_with_Bank) 2
91.	14.0	76.0	(Loans_Future) 2 => (Years_with_Bank) 2
81.	18.0	66.0	(Loans_Future) 1 => (Years_with_Bank) 2

The results show that the customers who are banking with the bank for less than 5 years have mainly a credit card only, apart from their S/B account (Rule 19). People who have found benefits of the bank are also the ones who have recommended the bank to 6-10 other people. Customers who have availed of a housing loan are the ones who have been with the bank for about 10 years. Also people who intend taking an education loan in the future tend to stick with the bank for 5-10 years.

It is found that customers who have been banking for more than 10 years are the ones who will not leave the bank easily.. The customers who are there with the bank for less than 5 years are the most delicate lot, in the sense that they would easily migrate to other banks the moment they are offered better service. So the bank has to continually upgrade its service for these customers. the customers banking for 5-10 years are a mixed lot, in the sense that it depends on the type of product they are availing of, i.e. if it is a SB AC chances are less that they will attrite, while a CA AC or CC customer will very easily turn to a different bank the moment they get enhanced service there.

Next-Sequential-Product Analysis: This is a 'very interesting to observe' analysis.

12.	20.0	79.0	(Total_Annual_Income)2=>(No_of_SB_AC)1
19.	13.0	66.0	(No_of_SB_AC)2=>(Total_Annual_Income) 3
48.	10.0	51.0	(No_of_CA_AC)1=>(Total_Annual_Income) 3
51.	11.0	62.0	(No_of_SB_AC) 1 => (No_of_CA_AC) 1
58.	15.0	52.0	(Loans_Taken) 6 => (Loans_Taken) 3
75.	18.0	73.0	(Loans_Taken) 3 => (Loans_Taken) 1
77.	11.0	63.0	(Loans_Taken) 3 => (Loans_Taken) 2
81.	13.0	72.0	(Loans_Taken) 1 => (Loans_Taken) 2
97.	16.0	82.0	(Loans_Taken) 2 => (No_of_FD_AC) 1

On analyzing the data collected it is found that people in the cluster 4 are the ones who mostly have a SB AC, education loan and credit card. These people once they come to cluster 3 mostly get married and take various personal loans for various purposes, consumer loans, etc. Customers in cluster 1 are the ones who mostly go in for housing loans, car loans, etc. The customers in clusters 1 and 2 are the ones who mostly have plans of taking loans in the near future. So they are the target ones for selling of loan products.
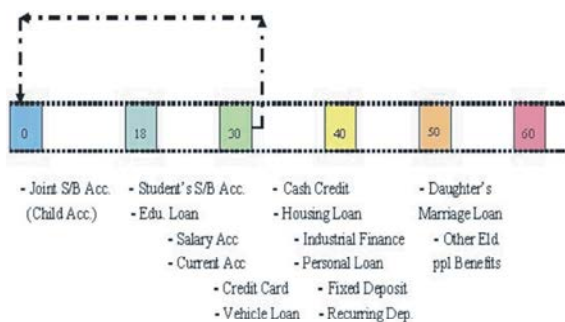

Fig. 1: Next Sequentiasl Product Analysis

Customer Satisfaction Analysis: Following is the output from the customer satisfaction analysis:

1.	16.0	74.0	(Recommend) 2=>(Any_other_Bank) 1
2.	14.0	66.0	(Recommend) 2=>Info_on_New_Prod) 1
3.	10.0	55.0	(Recommend) 1=>(Any_other_Bank) 1
4.	10.0	52.0	(Recommend) 1=>(Info_on_New_Prod) 1
17.	11.0	66.0	(Contact_Mode) 34=>(Info_on_New_Prod)1
41.	11.0	36.0	(Any.other.Bank)1 (Info_on_New_Prod) 1=>(Recommend)2
47.	13.0	63.0	(Age) 32-39=>(Info_on_New_Prod)1

It was found that depending on the branches the ambience and convenience of location of the branches varied. The variance in the response to ambience of branch was mostly restricted to 'good' (63.7%) and 'excellent' (23.05%) with 7.5% opting not to answer at all. As for waiting outside the ATM most customers have to wait rarely (82.1%).

Next Sequential Product Analysis: Based on the data collected the customer life-cycle has been designed. The banks also need to follow the customer life-cycle ardently. But prior to that they must also analyse the data from their customers and it may so happen that the bank may evolve with a life-cycle with some changes. Whatever may the case be, after finalizing on the customer life-cycle the banks must then keep track of the customers, Keeping track of the customers would mean the transactions that they do with the bank, any piece of information that the bank gets about the customer, etc. This information would get updated in the customer information data bases and using proper CRM tools the banks can analyse these databases to find patterns n their behaviour. Based on these patterns the banks need to reach the customers prior to their coming to the bank for the same purpose.

Churn Analysis: Churn analysis is a major task in banking and telecom sector. According to this analysis the system can classify a new customer as loyal or churn
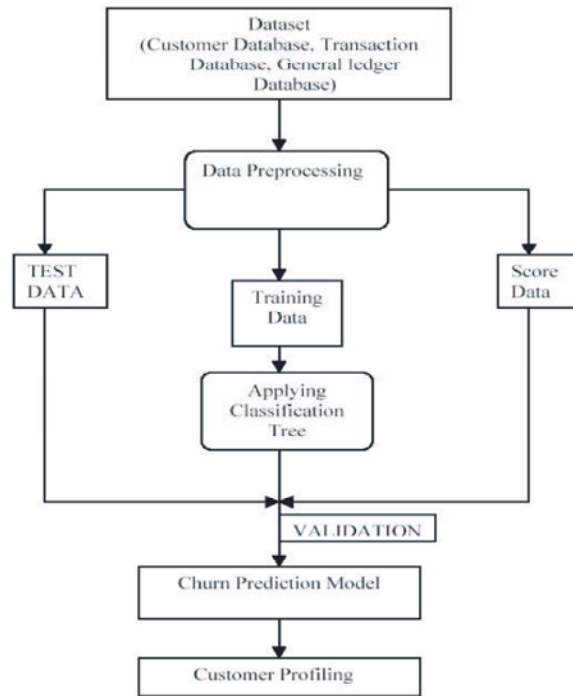
Fig. 2: Behavior analysis using CART

tendency. Behavioral profiles for defectors are identified using the if-then rules obtained by CART, which enables the bank to form an early warning system for churn prevention, considering that the customers exhibiting similar behavior to defectors are also likely to churn. It predicts the future churn for banking customers and can then be extended, thereby helping formulate intervention strategies based on churn prediction to reduce the lost revenue by increasing customer retention. It is expected that, with a better understanding of these characteristics, bank managers can develop a customized approach to customer retention activities within the context of their Customer Relationship Management efforts.

CONCLUSION

The study has first tried to analyse the need for CRM in the PSBs of India. The areas where the banks were lagging behind were identified. Further primary data was collected from the customers as to judge their perspective. This data was properly analysed using OLAP and data mining tools. Based on the analysis of this data certain recommendations were proposed that the banks may refer to in line with their own strategies. Further a Customer Life-Cycle was traced out based on the next-sequential-

product analysis. The Holistic frame work of CRM for Banks was also proposed. In the process Customer Value (CV) was also calculated.

REFERENCES

1. Dyche, J., and Dych, J. (2001). The CRM handbook: a business guide to customer relationship management. Reading, MA: Addison-Wesley.
2. Thomas, L.C., 2000. A survey of credit and behavioral scoring: forecasting financial risk of lending to consumers. International Journal of Forecasting, 16: 149-172.
3. Silberschatz, A. and A. Tuzhilin, 1996. What makes patterns interesting in knowledge discovery systems. IEEE Transactions on Knowledge and Data Engineering, 8(6): 970-974.
4. Blattberg, R., P. Peacock and S. Sen, 1976. Purchasing Strategies Across Product Categories, Journal of Consumer Research, 3: 143-154.
5. White paper: Increasing Customer Value: Improving CRM With Predictive Analytics Herb Edelstein, Two Crows Corporation
6. Handbook of consumer behavior (pp: 85-123). Englewood Cliffs, NJ:Prentice-Hall.
7. Gehrke, J., V. Ganti, R. Ramakrishnan and W.Y. Loh, 1999. BOAT – Optimistic decision tree construction, in Proc. ACM SIGMOD Int. Conf.Management of Data, Philadelphia, PA, pp: 169-180.
8. Thooyamani, K.P., V. Khanaa and R. Udayakumar, 2013. Blue tooth broad casting server, Middle-East Journal of Scientific Research, ISSN:1990-9233, 15(12): 1707-1712.
9. Thooyamani, K.P., V. Khanaa and R. Udayakumar, 2013. Weed control system of tea garden using GIS based database Management system, Middle-East Journal of Scientific Research, ISSN:1990-9233 15(12): 1702-1706.
10. Saravanan, T., V. Srinivasan and R. Udayakumar, 2013. A Approach for Visualization of Atherosclerosis in Coronary Artery, Middle-East Journal of Scientific Research, ISSN:1990-9233, 18(12): 1713-1717.
11. Saravanan, T., G. Saritha and R. Udayakumar, 2013. A Robust H-Infinity Two Degree of Freedom Control for Electro Magnetic Suspension System, Middle-East Journal of Scientific Research, ISSN:1990-9233, 18(12): 1827-1831.

12. Udayakumar, R., V. Khanaa and K.P. Kaliyamurthie, 2013. Performance Analysis of Resilient FTTH Architecture with Protection Mechanism, Indian Journal of Science and Technology, ISSN: 0974-6846, 6(6): 4737-4741.

13. Udayakumar, R., V. Khanaa and K.P. Kaliyamurthie, 2013. Optical Ring Architecture Performance Evaluation using ordinary receiver, Indian Journal of Science and Technology, ISSN: 0974-6846, 6(6): 4742-4747.