A Study on Customer Rentention using Predictive Data mining Techniques

Dr.V.P.Eswaramurthy[1], S.Induja[2]
[1] Assistant Professor, Department of Computer Science, Government Arts and Science College, Kumarapalayam, Tamilnadu, India.
[2]Assistant Professor, Department of Computer Science, Tiruppur Kumaran College For Women, Tirupur, Tamilnadu, India.

***Abstract*—** The customer is the soul or even essence of the organizational abilities. The age of globalization together with the competitors has totally changed the fundamental idea of merchandising, at the present merchandising is not tied to providing the services to the customer, and however the goal is to actively meet the hearts of the consumers. In the powerful industry circumstances where exactly companies are setting varying techniques now and then, customer retention is a crucial area to ponder upon, as customers usually churn from one organization to another. So prediction of the customer behaviour or even hence taking corrective actions before hand is the need of the hour. Data mining will help service sector like banking, insurance protection, and telecommunication to make critical business decisions. This paper throws light on the underlying technologies and the viewpoint applications of data mining in predicting the churn behaviour of the customers and hence paving way for better Customer Relationship Management.

Keywords—Data mining , Churning , CRM , Data Mining techniques.

I .INTRODUCTION

Customer Retention is a growing factor in today's ever-competitive commercialized world. Even more competition and higher standards made it more problematic for the organization to stand out from the crowd. Studies done at a world-wide and national levels write down the perceived importance of a few imperatives for business organization for their existence and growth.

Customer retention is foremost and effective method for the growth of the various organizations. Customer Relationship Management tools have developed and employed in order to strengthen customer acquisition and retention and to give necessary analytical steps such as predictive modelling and classification. Traditional churn. prediction methods have the advantage of being simple and even robust with respect to flaws in the input data, and they include severe limitations to describe the reasons for churn.

II. TYPES OF CHURNERS

As in fig.1, the two types of churners are voluntary and involuntary. Involuntary churners can be identified easily. It includes customers who are churned for misbehaviour, not using the phone properly,etc. Voluntary churners cannot be identified easily[13]. Without knowing, these customers will terminate their service.

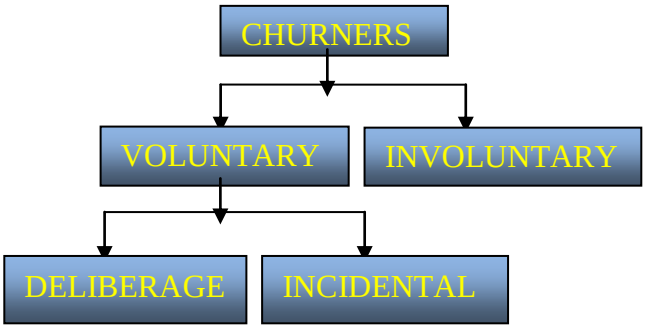

Fig 1: Taxonomy of Churn

III.DATAMINING

Data mining is the process of expedition and analysis, by automated or semi-automatic implies, of huge amounts of data in order to discover purposeful patterns and regulations. It can also be defined as the process of selecting, exploring and modelling large amounts of facts to uncover formerly unknown data forms for business benefit. Classification of large data sets is an important problem in data mining. The classification problem could be simply stated as follows. For a database with a number of records and for several classes such that each record belonging to one of the given classes, the problem of classification is to decide the class to which a certain record belongs. But there are much more to this than just simple classifying

Data mining has a lot of advantages and drawbacks depend on which particular industry used. The following are some advantages and drawbacks of data mining.

Advantages of Data Mining: Finance / Banking: Data mining supports for financial activities in banking including keep up the loan information of customers and credit card reporting.

Marketing/ Retail: It can help marketing organizations to set up the model based on the historical data such as online marketing, direct mail etc. Like Marketing data mining bring lot of benefits in retail enterprise.

Manufacturing: By using data mining techniques, manufacturers detect faulty devices and decide best control limits.

State organizations: Data mining helps government agency by digging and analyzing documents of financial agreements to build patterns that detect money laundering or illegal doing activity.

Drawbacks of Data Mining

Seclusion Issues: Due to the privacy issues, people are very much scared of their personal information that is collected and used in immoral way. Safety Issue: Security is such a big problem. In Business industry, how the information will be taken care is a big question. Misuse of information: The information may exploited by other persons and used for decision-making. It will cause serious problems.

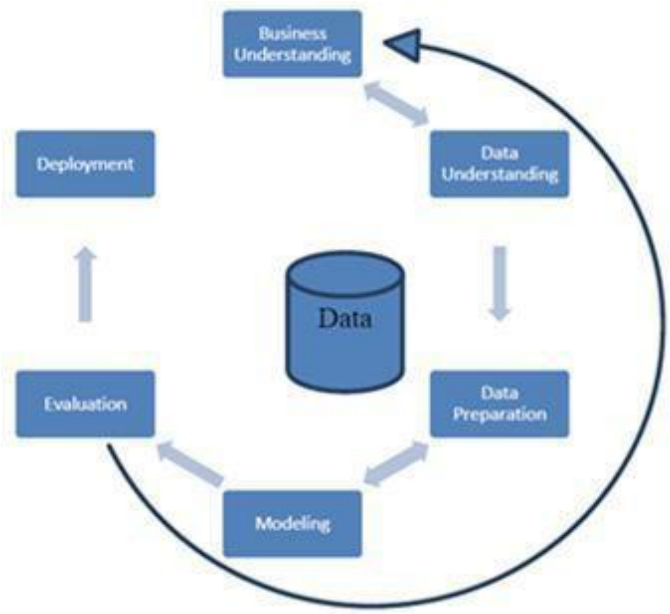

Fig 2: Phases of Data mining model

IV. CUSTOMER RELATIONSHIP MANAGEMENT

In the period of cut throat competition, the customer is considered as the Master. Data mining will be helpful for all the phases of Customer relationship cycle such as Customer Acquisition, Increasing Benefit of the Customer and Customer Retention [3].

CRM consists of three words Customer, Relationship, and Management [4]. Customers are those engines without which business are unable to run. Therefore the role of customer is very important [1].

Relationship: That attachment which associations have to create with their customers is called relationship. Administration: It does involve all those actions which have to prepare organize motivate and manage dealings with customers [1].

Definition for CRM:

It is defined as "the process of acquiring, retaining and growing profitable customer which requires a clear focus on service attribute that represent value to the customer and creates loyalty"[2]. CRM is relationship marketing, which aims are improving the long-term profitability of customers through moving away from product-centric marketing to customer-centric[10] . CRM is known as a chance to increase gain, magnetizing and retaining economically useful consumers through removing "economically invaluable" ones [5].

According [2], CRM consists of four dimensions: Consumer Classification, Consumer Attraction, Customer Retention and Customer Progression. All of these four measures can be regarded as blocked loops. Customer Identification: CRM starts with consumer identification which is often known as customer acquisition. This phase involves targeting the population who are most likely to be a consumer or profitable to the organization. Components of consumer classification include target consumer analysis and consumer segmentation. Consumer Elegance: This is the following phase after the customer identification. It fascinates the objective customers. After identifying the segments of potential of consumers, organizations could direct effort and resources into attracting the targeting customer.

Customer Retention: This is the central concern for CRM. Client satisfaction which denotes the review of customer expectations, it is the essential problem for retaining customer's. Elements of customer retention comprises of one-to-one marketing, honesty programs and disorder management. Customer development: This involves consistent enlargement of financial transaction cost and unique consumer potential earning. Elements of customer progression include customer lifetime value analysis, up/cross selling and market basket analysis. Market analysis goal is to maximizing customer transactions and reveal trends in customer's behaviours..

Applications of data mining for CRM

The applications of data mining are split into following various kinds. Some of the few applications areas and also highlighted how data mining tools have to be developed for

such type of application.
Financial data analysis: Financial data collected in the banking and financial industries are often really complete, consistent and of high quality which features systematic data analysis and data mining. Some examples cases such as Design and construction of data warehouses for multidimensional data assessment and data mining, Loan payment forecast and consumer credit policy analysis, Identification and clustering of customers of aiming marketing, Detection of money laundering and other financial crimes.

Retail Industry: It will be the major utilization area of data mining because to identify consumer buying behaviour, achieve good customer preservation. Retails data mining can help pinpoint consumer buying behaviours and obtain better customer retention and satisfaction. A few examples of retail industry are Model and construction of data warehouses based on the benefits of data mining,, Multidimensional analysis of income, consumers, products, period and area, Research of the efficiency of promotional strategies, Consumer retention.

Health Care and Insurance Industry: The development of insurance industry fully depends on the potential of transforming data into the expertise information and facts or intelligence about customer's challengers and its markets. Data mining also applied in statements analysis such as identify which medical treatments are recognized together. The techniques those are useful for scams detection in insurance. In insurance, four groups just like home insurance, term life, motor insurance, and medical insurance. Among the four, the vehicle and health insurance have got much more fraudulence problems. The data mining techniques which can be more helpful for detecting the fraudulence in the insurance sector.

Communication technology industry: Communication technology industry is rapidly expanding and strongly competitive. It creates a demand from customer, data mining in order to understand the business, recognize the communications routines and improve the good quality of services. The circumstances for which data mining may strengthen telecommunications services as such as Multidimensional analysis of telecommunications data, Liable pattern analysis and the identification of uncommon patterns, Mobile communications, use of visualization tools in communications data analysis.

V.PREDICTIVE DATAMINING TECHNIQUES

It may also be defined as the method of choosing, exploring and modelling large amounts of data files to find out formerly not known data patterns for organization benefit. The classification of huge data collections is an essential issue in data mining. The classification issue can be stated simply as follows. For a database with numerous records and for several classes such that each one document belongs to one of the given classes, the problem of classification is to decide the class to which a given document belongs. But there is however much more to this than simply very simple classifying.

There are many predictive classification techniques namely nearby neighbour technique, decision tree technique, linear discriminate technique, naïve bayes technique, etc. In this paper, decision tree technique is used. The specific specifications that should be taken into account while designing any decision tree construction algorithms for data mining are that
a. The technique should be efficient to be able to handle a very large database object.
b. The technique should be able to manage categorical attributes.

- Decision Tree
- Regression Analysis
- Fuzzy logic
- Neural Networks

VI. REVIEW OF LITERATURE

M. PurnaChander[6] highlighted that obtaining new customers is more costlier process than retaining existing customers. They highlighted on modelling churn behaviour of bank customers in Indian circumstance. Many characteristics of customers such as demographic details, psychographic, product purchase details, customer perception details are vital in modelling the churn habits of bank customers. In Indian banking system, the banks haven't arranged their data earlier as per the specifications for prediction of their churn behaviour, so they highlighted the method of converting real customer data into purposeful data that meets modelling churn behaviour. Then, this purposeful data can be converted into practical knowledge; predictive data mining techniques are used.

U. Devi Prasad[7] emphasized that the use of data mining tools topredict the churn behaviour in the banking sector.The unprocessed data from collected from banks may be transformed into meaningful and useful facts and then into practical knowledge in which predictive data mining techniques are used. They had experimented with 2 classification techniques namely CART, and C 5.0 . the prediction success cost of Churn class by CART is quite high but C 5.0 had shown bad results in predicting churn customers. Although, the prediction results rate of Active class by C 5.0 is more powerful than the other technique. They have considered dynamic time period, which differs for every customer. In this analysis, they have experimented with classification techniques namely CART and C 5.0 on 1.484 samples of bank customers in which 1.163 were active customers and

311 were churn customers. While CART gives 95.1%classification rate on training data and 91.22 on test data, C5.0 yielded 69.3%t classification rate on training data and 68.9 per cent on test data. Thus, a model with prediction good results cost of Churn class to be chosen for reaping larger profits. The research forecasts the churners of of banking customers that can be monitored ,by following involvement strategies depending on the churn forecast by lessening lost income by raising customer preservation .It was predicted that, with a better understanding of these characteristics, bank managers can build up a custom-made method to customer retention actions within the Customer Relationship Management ventures.

Indranil Bose[8] brings out two-stage hybrid models to group unsupervised learning technique with supervised learning technique. It developed a model for the Customer churn Prediction. The most important thing in customer churn management is segregating churners from non-churners. Decision tree model are the most popular model in the churn prediction. It used various variables for clustering and analyses different method of hybridization for getting the results of clustering in order to form supervised learning models for churn prediction.

S. Balaji [9] used Naive Bayesian Classification algorithm for customer classification and to predict churners that are churned on Life Insurance sector. He also used Naïve Bayes classification to classify the customers from larger dataset. It also analyses the issues of using data mining technology for predicting the customer habits. In this analysis, they had tested 10.000 sample of Life Insurance of customers, the unprocessed data can be converted into useful information and then into knowledge for which they had used predictive data mining techniques .Posterior classification process applied for the data in this paper. It proved that the naïve bayes classifier is much better than other classifier for conducting the policy preferences towards the customers. This helps us to raise the income of the organization

Narender Kumar[10] makes use of K-means algorithm to build a model to find out the relationship in a customer database. Cluster analysis will find out the group of persons belongs which criteria. The customer data from LIC have taken in account for the analysis. K-means is the suitable technique for cluster analysis. It will increase the profit of the organization. Clustering optimization method is used to find the appropriate or local optimal solution.

K.Chitra [11] used Regression tree for yielding better results. In this paper, they had used CART which is one of the most popular methods for building decision trees .CART, by splitting the records at each node, builds a binary decision tree, according to a function of single attribute. They had build a model for predicting churners in Banking Sector. In this paper, they had analysed more than 1000 of records and used various attributes for customer classification. They had given a solution for the prediction of customers using CART and future enhancement.

P. Isakki alias Devi [12] had developed a method to design retail promotions, observed in the same groups of customers. It used the Clustering and association rule to find out the identity of the customer behaviour. It can also easily forecast the sales. The customer with similar behaviour is grouped by clustering techniques such as K-means method. For each cluster an association rule (Apriori algorithm) was used to identify the products bought by the customers. Relationship and knowledge of the databases are discovered by using Association rules. Apriori algorithm, the most well known association rule mining algorithm as it easily found out the frequent item from the database was proved in this paper. WEKA, the open source data mining tool was used to perform the analysis . It also states that the analysis of customer behaviour will improve the overall performance of the organization and more customer satisfaction.

VII .CONCLUSION

This survey paper concludes that the data mining techniques is very important for the customer relationship management. In this paper, we have shown that various data mining techniques can be grouped for the enhancement of CRM . Data mining will have a major impact of CRM and will face and present challenges for future research. By judicially selecting, the data mining techniques and their proper usage may prove to become a boon for the organization and they will be a chance, to avoid churning of customers that leads to very long term relationship of the customers with the enterprises.This work gives a study of different data mining techniques that are used in the prediction of churners. The next stage of the research will involved in performing a deeper analysis into the customer database to build a churn prediction model. For Classification and Clustering the hybridization techniques can be used in future to solve different CRM problems.

REFERENCES

[1] S.Janakiraman,K.Umamaheswari, "A Survey of Data mining techniques for Customer Relationship Management",International journal of Engineering, Business and Enterprise application.

[2] E.W.T. Ngai, Li Xiu, D.C.K. Chau, "Application of Data mining technique in customer relationship management: Al literature review and classification", Expert Systems with Applications, Vol.36, 2009, pp. 2592-2602.

[3] Kazi Imran, Dr. Qazi Baseer Ahmed, " Use of Data Mining in Banking", International Journal of Engineering Research and Applications,Vol.2, No.2, April 2012, pp. 738-742.

[4] Rajni Arora, "Customer Relationship Management", International Journal of Research in IT & Management, Vol. 3, Issue 8, August 2013, pp. 48-57.

[5] Nastaran Mohammadhossein, Dr. Nor. Hidayati Zakaria, ," CRM Benefits for Customers: Literature Review (2005 –

2012) " , International Journal of Engineering Research and Applications ,Vol. 2, Issue 6, December 2012, pp.1578-1586

[6] M. Purna Chandar, Arijit Laha and P. Radha Krishna - Modeling churn behavior of bank customers using predictive data mining techniques", National Conference on Soft Computing Techniques for Engineering Applications,.

[7] Dr. U. Devi Prasad, S. Madhavi, "Prediction of churn behavior of Bank Customers", Business Intelligence journal, Vol 5, No. 1, January 2012

[8] Indranil Bose, Xi Chen, "Hybrid models using Unsupervised Clustering for Prediction of Customer Churn", Proceedings of the International MultiConference of Engineers and Computer Scientists, Vol. 1, March 18-20, 2009, Hong Kong.

[9] S. Balaji, S.K. Srinivasta, "Naïve Bayes Classification approach for Mining Life insurance Databases for Effective Prediction of Customer Preferences over Life Insurance Products", International Journal of Computer Applications, Vol.51, No. 3,2012.

[10]] Narander Kumar, Vishal Verma, Vipin Saxena, " Cluster Analysis in Data Mining using K-Means Method", International Journal of Computer Applications , Vol. 76, No. 12, August 2013, pp. 11-14.

[11] K.Chitra,B.Subhashini, "Customer Rentention in Banking Sector using Predictive Data mining Techniques" ICIT 2011 The 5th International Conference on Information Technology.

[12] P. Issakki Alias Devi, S.P. Rajagopalan, "Analysis of Customer Behavior using Clustering and Association Rules", International Journal of Computer Applications, Vol. 43, No.23, April 2012, pp.19-27 .

[13] Vivek Bhambri, "Data Mining as a Tool to Predict Churn Behavior of Customers", International journal of Computer & Organizatio Trends , Vol. 2, Issue. 3, 2012, pp. 85-88.