



## Data Mining model for Insurance Trade in CRM System

S.Balaji<sup>1</sup>, Dr.S.K.Srivatsa<sup>2</sup><sup>1</sup> Research Scholar, Vels University<sup>2</sup> Senior Professor, St. Joseph Eng. College Chennai-600100

Email: srisaibalaji@rediffmail.com

**Abstract:-** Data mining is used to extract meaningful information and to develop significant relationships among variables stored in large data set/ data warehouse. In this paper data mining technique named k-means clustering is applied to analyze customer's preference over Insurance products. Here K-means clustering method is used to discover knowledge that come from CRM system. This study will help the to identify the customer's preference over Insurance products and improve the product selling strategies.

**Keywords:-** Data Mining, Clustering, K-Means.

### 1. Introduction

Data Mining is a process of extracting previously unknown, valid, potential useful and hidden patterns from large data sets (Connolly, 1999). As the amount of data stored in CRM databases is increasing rapidly. In order to get required benefits from such large data and to find hidden relationships between variables using different data mining techniques developed and used (Han and Kamber, 2006). Clustering and decision tree are most widely used techniques for future prediction. The main goal of clustering is to partition customers into homogeneous groups according to their characteristics and abilities (Kifaya, 2009). These applications can help both advisors and customers to enhance their decision making over the Insurance products.

This study aims to analyze how different factors effect a customer's preference over Insurance products using k-means and decision tree in CRM System. Decision tree analysis is a popular data mining technique that can be used to explain the interdependencies among different variables such as financial status and product preference. Clustering is one of the basic techniques often used in analyzing data sets. This study makes use of cluster analysis to segment customers into groups according to their characteristics. The remaining parts of the paper are organized as follow: section 2 described the related work in CRM Data Mining. Section 3 provides a general description of the data we used. Section 4 described the process stage of data used. Section 5 reports our experimental analysis of data mining methods applied on CRM data set. Finally, we conclude this paper with an outlook for future work.

### 2. Related Work

(Kifaya, 2009) K-means clustering is a widely used method that is easy and quite simple to understand. Cluster analysis describes the similarity between different cases by calculating the distance. These cases are divided into different clusters due to their similarity. In {Peter C. Verhoef, Bas Donkers, 2001} gave Predicting customer potential value an application in the insurance industry. In [Ruey-Shun Chen, Ruey-Chyi Wu and J. Y. Chen, 2005] gave Data Mining application in customer relationship management. In (Han and Kamber, 2006) explained that k-means is a well known clustering algorithm tends to uncover relations among variables already presented in dataset.

### 3. Data Mining

Data mining software allow the users to analyze data from different dimensions categorize it and a summarized the relationships, identified during the mining process (Han and Kamber, 2006). Different data mining techniques are used in various fields of life such as medicine, statistical analysis, engineering, education, banking, marketing, sale, etc (MacLennan., 2005). Cluster analysis used to segment a large set of data into subsets called clusters. Each cluster is a collection of data objects that are similar to one another are placed within the same cluster but are dissimilar to objects in other clusters. (Behrouz.et.al., 2003, Dongsong, 2004). Data mining techniques are used to operate on large volumes of data to discover hidden patterns and relationships helpful in decision making (Connolly, 1999).

### 3.1. Data mining in Customer Relationship Management System (CRM):-

Customer Relationship Management (CRM) systems are adopted by the organisations in order to achieve success in the business and also to formulate business strategies, which can be formulated based on the predictions given by the data mining tools. Data mining is a new technology that helps businesses to predict future trends and behaviours, allowing them to make proactive, knowledge-driven decisions. When data mining tools and techniques are applied on the data warehouse based on customer records, they search for the hidden patterns and trends.

Every company aspires to implement methods within their business through which they can attract new customers and retain existing customers. The methods to do this have changed noticeably over time through the ever evolving advent of technology. This technology has excelled primarily due to the existence of the world-wide web and the various methods of relaying information through the numerous wired and wireless networks available. Due to this, CRM has become an intriguing area of research for various industrialists and academics alike. CRM allows a particular company to build and maintain customer relationships in such a way that it can use every avenue of its business to entice the customer and therefore increase its profit holdings (Zablah et al 2004).

### 4. Proposed Model

In a Policy Holder's database the attitude of the customer is determined by their financial status as well as need assessment. Financial status is made on the basis of age, Occupation, education, Income, Marital status, career position and external support. While at the same time need assessment of a customer based on family need, occupational need and occasional need.

The proposed model makes prediction about traditional policy and Unit linked policy preference of customers based on financial status as well as need assessment. The proposed model also deals with policy holders in entrance ratio of customers in a particular policy and renewal status of the customers after holding policy. The Model was developed using SQL queries available in Visual Program.

#### 4.1. Application

In this study, data gathered from customers was analyzed using a data mining technique namely k-means clustering. In order to apply this technique following steps were performed in sequence:

##### 4.1.2. Data Set

The data set used in this study was obtained from Insurance regulatory of India in 2010. includes collective premiums by LIC of India and the 22 private life insurers.

##### 4.1.3. Database

The database management system used in this study was Microsoft SQL server 2005. This software was used because it was compatible and efficient to use with the database management system i.e. relational database and the other reason was that the data was maintained in this database prior to the study.

##### 4.1.4. Application Software

The programming environment used for application was Visual studio 2005 for building data mining model. It was suitable for development of mining model and was compatible with SQL Server 2005, in which data was maintained/ stored.

#### 4.2. Data mining Process

Data mining process consist of following steps:

##### 4.2.1. Preparations

In this step data stored in different tables was joined in a single table after joining process errors were removed.

##### 4.2.2. Data Selection and Transformation

In this step we determined the fields of study used for analysis. Data is inform of yes/no is transformed in form of 1/0.

##### 4.2.3. Implementation of Mining Model

In this step, k-means clustering algorithm was applied to the processed data to get valuable information. K-means is an old and most widely used clustering algorithm developed by MacQueen in 1967 (Erdogan and Timor 2005).

#### Algorithm 1.

##### Basic K-Means algortihm

1. select K points as the intial centroids
2. repeat
3. Form K clusters by assigning all points to the closest centroid
4. Recompute the centroid of each cluster
5. until The centroids donot change

#### 4.3. Results and Discussion

The model produced following results

In the existing insurance market, where numerous insurance companies are selling a host of policies, buyers are spoilt for choices.

We grouped the customers regarding their policy classification in several ways 3 of

Which are:

1. Assign policy labels as option for a traditional insurance or a market-linked plan.
2. Group the policy holders as Single premium policy holder or a regular premium.
3. Categorize the customer policy with one of two class labels Group policy or Individual policy holders.

	Premium collection Value in Rs.in Crores	No .of Policies Sold	Percentage as Per value	Percentage as per Nos.
Group Policies	25386.66	17,734	32.97	0.06
Individual Policies	51603.22	2,63,34,233	67.03	99.94
Non-Single/Regular Premium Policies	38215.68	22,54,6028	49.64	85.55

	Premium collection Value in Rs.in Crores	No .of Policies Sold	Percentage as Per value	Percentage as per Nos.
Non-Traditional policies/Market Linked Policies	45215.68	39,64,050.802	58.73	15.04
Traditional	31774.20	2,23,87,916.198	41.27	84.96

Table 1:-

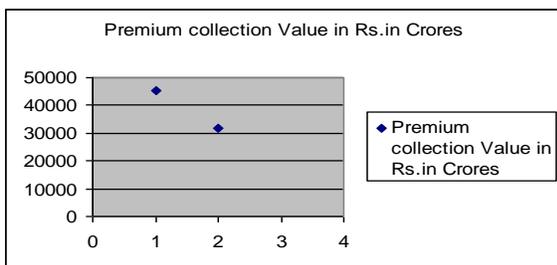
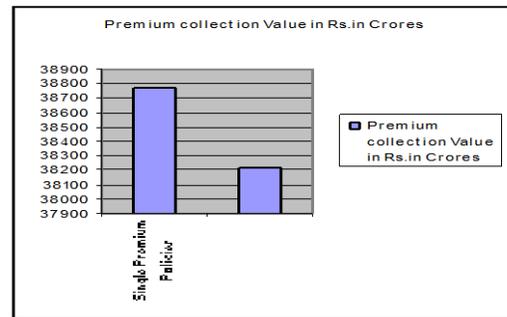
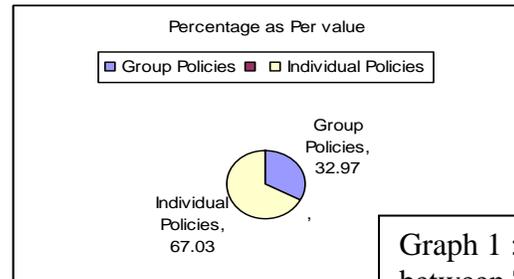


Table 2:-

	Premium collection Value in Rs.in Crores	No .of Policies Sold	Percentage as Per value	Percentage as per nos.
Single Premium Policies	38774.20	38,05,939	50.36	14.45



**Graph 2:-Premium collection value of Single and Regular Premium**



**Graph 3:-Percentage of policy holders having Group and Individual Policies**

**Graph 1 :-Relationship between Traditional and Non-Traditional Policies-Premium Collection**

Findings:-As expected most insurance seekers are looking for ULIPs and this is very much in line with the recent product distribution declarations by most insurers. Term plans form the next big chunk. A qualitative input, will spell out the opportunity here very clearly- most customers who wanted ULIPs were not sure of the product features and how to evaluate ULIPs.

## 5. Conclusion:-

In this study we make use of data mining process in a insurance database using k-means clustering algorithm to predict customer's product preferences. We hope that the information generated after the implementation of data mining technique may be helpful for advisors as well as for customers. This work may improve advisor's performance, reduce failing ratio by taking appropriate steps at right time to improve the quantity of customers . Many leading insurance companies are making an effort to move away from the product-oriented architectures of the past and toward a customer-centric focus to better serve their customers. Customer-centric marketing is accomplished by integrating data mining and campaign management. By assessing predicted customer behavior, marketing professionals can further refine the marketing campaigns. The key to gaining a competitive advantage in the

insurance industry is found in analyzing this data and getting a greater insight into their business.

## References:-

- [1] Peter C. Verhoef , Bas Donkers, Predicting customer potential value an application in the insurance industry, Decision Support System[M], 32, 2001, 189-199.
- [2] Janny.C.Hoekstra, Eelko.K.R.E.Huizingh, The Lifetime Value Concept in Customer-Based Marketing, Journal of Market Focused Management [M], 3, 1999, 257-274.
- [3] (US) Alex Berson, Stephen Smith, Kurt Thearling, Translated by Qi He, Yan Zheng, Establishing Data Mining Application in orienting CRM, (2001).
- [4] Yanjie Lv, Tao Yin, Qi Wang, Customer Relationship Management and its Subject Analyses[