Using Data Mining for Assessing Customer Value in Insurance by Neural Network

Seyed Reza Oroumyeh[1], Saeid Soleymani Vayghan [2]and Fardin Behboudi Vayghan [3]

Artificial Intelligence, Computer-Software, Computer-Software

Khaje Nasir Toosi University(kntu), Azad University of Shabestar, Azad University of Shabestar

ABSTRACT

One way to increase income for most insurance companies is selling more products to the best customers. On the other hand, the data from the insurance company for mass records are kept and analysed, a database of individual behaviour analysis for insurance clients is difficult and the development of new products and selling insurance to individuals and organizations is important for insurance companies. Data from past clients on the insurance company data bases are available and are able to make decisions about new products offered to customers. In other words, the past behaviour of the client and the products used could be used for the value of the costumer.

KEYWORDS

Customer value, Data mining, Insurance, Packaging sector, Associative rules,Neural Network

1. Introduction

Most important purpose of marketing in the present age is to know the needs of the costumer and then to try and satisfy those needs. a database of past clients and customers who already are related participated , Richest source for data mining , Artificial intelligence algorithms and knowledge discovery is and The data comes from customers that can be used to explore the behavior of customers. This type of behavioural data can be used to explore the potential value of the customer, Prediction of their future needs, prediction of Customers credit risk and risk abandoning the product used. In this study, firstly presenting a customer clustering in order and then discover the rules contained in each cluster will then provide the solution for the analysis of data related to vehicle insurance, one of the country's insurance companies will work great. The development of new products and offer more products to current customers is one of the objectives of insurance companies. In the marketing strategy the behaviour in dealing with customers is known as Retention.While customers who use products and services from the company have to be treated in a different way [2].In part 2 we'll talk about its data collection then in section 3 the concepts of Cross-selling will be raised, and in the next section will talk about the use of data mining concepts and at the end the results obtained from the analysis of data mining will be raised.

2. Collection of Data

The data set under consideration include approximately 65,000 third-party insurance record and the vehicle hull insurance is one of the country's insurance agents during the past 6 years. The information was stored as database of the various and scatters. Figure 1 indicates this fact.[1]

[1] S.R.Oroumyeh Artificial Intelligence, Master Student of khaje nasir toosi (kntu),Tehran,Iran(email:SR.oroumyeh@kntu.ac.ir)

[2] S.S.Vayghan Computer-Software, Azad University of Shabestar,Shabestar,Tabriz(email:SSV_1368@yahoo.com)

[3] F.B.Vayghan Computer-Software, Azad University of Shabestar,Shabestar,Tabriz(email:F_behboudi@hotmail.com)

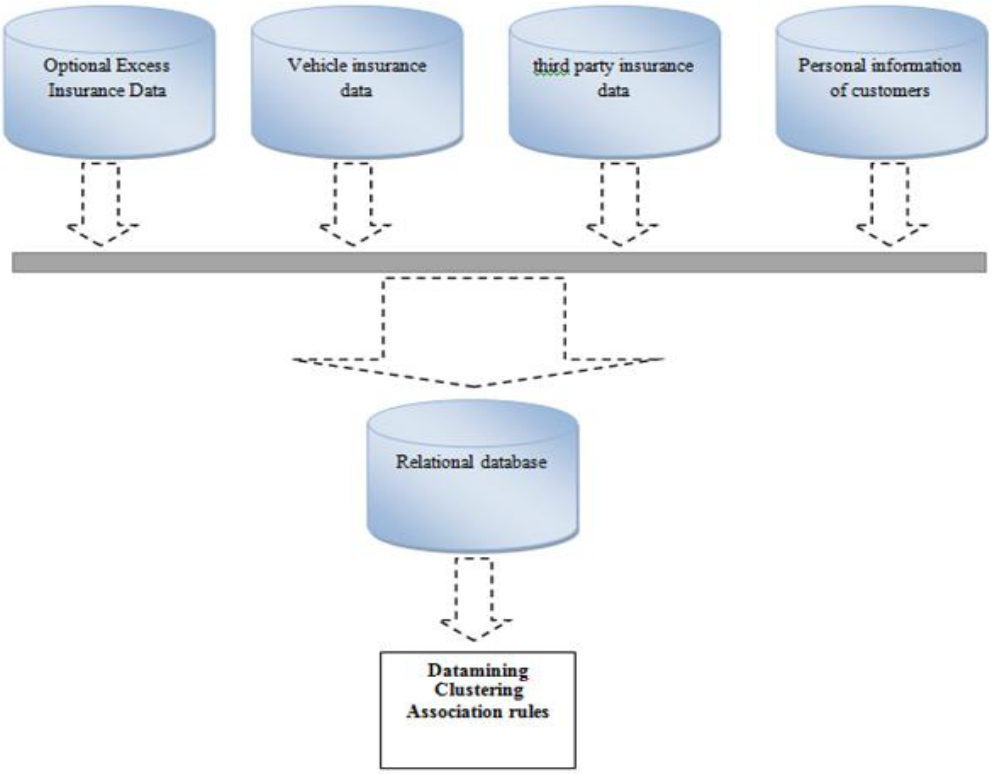

Figure. 1. Data integration Insurance Company

3. Costumer Value

Customer segmentation is one of the major applications of data mining. The objective of this Division is to fit up goods and services and marketing messages with each customer.
In the traditional method of dividing customers mainly from the statistical-based approaches were used to profile. The problem that these methods found was that there was a possibility of two people who, for example, in terms of age were the same but didn't have the same values and behavior. Data mining, this problem is solved as well.

A. Finding Categories based on Behavioural and Psychological Factors

One of the methods of the behavioural categories found in the use of clustering methods were not-redirected. Of data mining algorithms that are used mostly in this context can be Mention to k-means and neural networking. For example, in [3] ways to divide customers in online stores In terms of psychological methods and their willingness to buy, A method based on SOM and k-Means algorithm is presented

B. Find the Invoice Value According to Categories

Customer segmentation based on the value of each customer (this value to the benefit of the customer or company ever comes that he intends in the future due to customer loyalty will be on the company's existence) is one of the most common methods of segmentation.
Costumers have two types of value [8]:

1) Direct value: worth the money the customer
during customer lifecycle and long-term value for the company's needs.

2) The value of non-monetary value: indirectly

to such as customer loyalty and credibility.

Of the other work done in this area to techniques for discovering the patterns within each category and the analysis of the business environment, it can be named. For example, for customers to discover credit card transaction with the categories above can be a good example.

C. Determine Customer Value

Determine customer value is quite complex and difficult and the method of data mining plays an important role in this matter. Methods of determining customer value are primarily based on financial definitions. A very simple method for the calculation of customer value can thus be expressed (Barry and liaof):
Revenue from the customer-customer maintenance cost But the question is how much profit can be attributed to a customer? Do we consider their monthly purchase? How indirect values such as customer loyalty for our calculation? Calculate the cost of calculating profit is also more complex; the business may have different fees according to each customer. For example, it is possible that two costumer have different space to the commodity distribution warehouse space so the cost of shipment to a customer will be more than another; do this can the value of a customer be less than another?

4. Customer Relationship Management

The early years 1980, CRM importance in the marketing topics won. Although providing a comprehensive definition of CRM can be difficult it can be to face the overall CRM strategy overview that includes process absorption, Maintenance and accompany with selected customers, and it is going to create value for the company and the customer takes [4].The research of in line with the use of data mining in customer relationship management.

A. Customer Segmentation

The process of dividing the business customers to redeem the following distinctive, with a meaning and a form according to their specifications and features. Division make it possible for business provides its clients better understand the different marketing strategies for each group in the comments, and custom products tailored to each proposal.
The first customers are divided by Vernel Smith in 1956, was raised by a certain standard of customer supply Division [5]. Various research has shown that customers of different aspects have different shopping habits like together and alternately purchase or use of the services. In the marketing this would be remembered as the incongruity of the basic concepts in marketing is to find your target clients, market segmentation and marketing wisdom goes to work [6]. The clustering method for putting customers in the same category with the characteristics of the application.
We have three methods of clustering:

- The Non-overlapping clustering: each customer only belongs to a category.

- overlapping clustering: a client can belong to multiple groups.
- Fuzzy methods: the probability to distinct groups of belonging.

The Gartner Group, a consulting company known in the United States believes that customer segmentation means that the management of a broad vision for the organization that provides the ability to communicate with customers for the company provides and increases the rate of return customers [7].

B. Segmentation Based on Costumer's Value

Costumers are divided Depending on the value that they have for the organization.This Division is important because through it we can understand who the most valuable customers are. For example, a division based on customer value can be as follows:

- Golden batch: 20% of customers with the value
- Silver batch: 30% of customers with high value,but in the second row than the Golden batch
- Bronze batch: customers with the lowest value

The current active customers can be divided into the following categories and its use as a future decisions [7] Figure2 signifies Division provided that VIP customers into 4 categories.

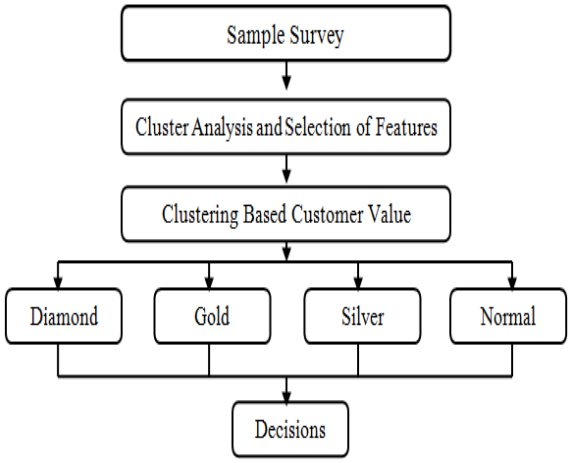

Figure 2: Classification of customer value

The segmentation method in data mining, intentionally so, based on the clustering method and category. Before the transaction data clustering customers should use the model to analyze customer behavior based on any payment. Combining clustering and these models we can provide a tremendous improvement in the more accurate customer segmentation for continuing to achieve CRM.

5. Data mining

Data mining , extraction knowledge of data available. Data mining with the help collection of methods statistical analysis and modeling, can be hidden patterns and relationships available in the data base of diagnosis [8]. So far various tools and techniques for information processing build results has been developed that the of manufacturing data storage information and establishing databases easily.Today organizations are able to receive a lot of information about their business with a low cost and using the methods of data mining could give the organization a remarkable value.

A. Association Rules

Finding association rules, is one of the most important tasks data mining and So far, much research has been done in this area. A common example is to find the set of goods that are frequently purchased together. Despite the large volume of data that be collected and stored everyday , For many businesses today to find such patterns is very valuable and can be in Cross-Selling, Catalog Design and Be used to analyze customer behavior.[9]
Suppose $I=\{i_1,i_2,i_3,\ldots,i_m\}$ as the set of items and D is a dataset containing transactions where each transaction T is a set of items, Such that $T \subseteq I$. An association rule is represented as an entailment $A \rightarrow B$,such that $A \subset I$, $B \subset I$. support (A,D) , represents the ratio of the transactions of D that contains a set of items X. Support, represents the statistical validity of D and whatever the this value is greater, The D credit is higher. support (A,D) , represents the ratio of the transactions of D that AUB is included. The standard evaluation for the rule $A \rightarrow B$ for is Confidence.
These values are defined as follows:

(1) Support((A→B) = P(AUB)
(2) Confidence((A→B) = P(B | A)= Support(AUB) / Support(A)

Given a set of transactions, Association rules with the least amount of production Obtaining Support, and Confidence is done by us, Needs to be done.

Thus resulting rules must include the following two conditions :

(3) Support(AUB , D)≥MinSup
(4) Confidence((A→B) ≥ MinConf

TABLE 1.
Percentage of correctly clustered

Partition	1_Training		2_Testing	
Correct	48.312	96.23%	12.081	96.13%
Wrong	4.070	3.77%	1.032	3.87%
Total	52.382		13.113	

B. MLP Clustering and Comparison

Clustering, segmentation similar patterns to found in data to clusters non-common. Here we have grouped them into four categories and using Multilayer Perceptron Neural Network classification work carried out on the data. In this approach, we used two hidden layers of MLP , The first hidden layer of 20 neurons and in the second hidden layer of 15 neurons were used. In this paper we better results obtain than the previous article that the K-Means clustering with the way done. Here's our 92.23 percent were correctly classified. this is method better that of K-Means.(table1)

6. Methods Used

A. Customer Segmentation

Insurance companies tend to have a thorough understanding of their behavior patterns and values and to use it as a competitive advantage. They should be on the behavior and demands of their customers and value their work. Therefore, the identification of groups that have the same characteristics makes possible for companies,that Group your target customers better manage and also provide customized products to suit each group.Therefore we of clementine software and using neural network clustering to clustering insurance customers paid.

Clustering results are as follows:

1) Diamond Cluster

The average age of the customers clusters are above 38 years. approximately These customers have high-priced cars and Most this vehicles manufactured after 2010. jobs Most of these clusters,are a engineering and medical. in these clusters customers purchasing power high and low credit risk and There are no damage history. These customers can be considered as the most valuable category. In these clusters customers purchasing power normal and low credit risk.

2) Gold clusters

The customers of this cluster don't use insurance and There are no damage history. The average age of the customers clusters are above 25 years And most have a private car and Most this vehicles manufactured after 2009.

3) Silver cluster

The average age of the customers clusters are above 40 years and all of their certified level1 are.most of these customers are driving jobs and truck insurance as well.No damage is over 90% of these customers and the rest is down the amount of damages.

4) Normal cluster

Accident rates are 65 percent of the customers this cluster. The average age of the customers clusters are above 31 years and most have a private car by The average price of domestic production are. jobs Most of these clusters, is employee . Most this vehicles manufactured between 2007 and 2009. In these clusters customers purchasing power lower than other clusters.

B. Association Rules Applied in Clusters

After clustering customers, We arrived to 4 clusters using neural networks. Then we examine the characteristics of each cluster; within each cluster ,follow relation to each of the attributes in each cluster, Thus, association rules and Apriori methods were used. Using the results of each cluster about buying life insurance can be predicted. Diamond Cluster and Cluster Gold, The best option for the sale of insurance products such as life insurance and supplemental insurance. In Figure 3, the Apriori algorithm is applied on the clusters diamonds.

7. Comparison

In previous research, we use the methods to classify K-Means. But here Using neural network to achieve a high percentage of classification accuracy. We have divided this article into 4 categories and we found a group that has the least damage to the insured. We Recommend , sale of insurance products in this category with lowest prices. Actually, If the insurance company to your products with lower price offer diamond and gold, Will further benefit.

8. Conclusions

In this paper a data mining based approach to assess the value we offer customers insurance. To this end of clustering algorithm NN, we use ,for insurance customers segmentation for vehicles .The analysis performed on the resulting clusters, we find that, Valuable clusters to offer insurance products diamond cluster and gold cluster. Customers located in this of the clusters according to past history insurance, investment risk is less and are higher than other clusters of purchasing power and probability purchase of insurance products by this the clusters ismuch higher than other clusters. In order to find the relationship between customer characteristics, association rules and Apriori methods were used and the relationship between the characteristics of all the clusters found. Insurance companies utilizing of the rules can be more careful in selecting their own customers.

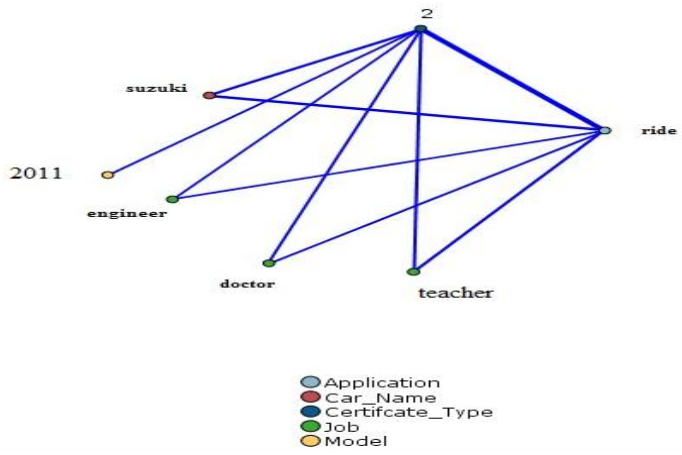

Figure 3: Apriori algorithm is applied to a cluster of diamonds

References:

[1] S. H. Mohammadi , M. Karamipour , S. R. Oroumyeh, " Using data mining for assessing Customer Value in Insurance", Iran Data Mining Conference (*IDMC*) , 2012.

[2] M. J. A. Berry, G. S. Linoff, "Data mining techniques : for marketing, sales, and customer relationship management" third ed,Wiley Publishing, 2012.

[3] T. Hong, E. Kim, "Segmenting customers in online stores based on factors that affect the customer's intention to purchase" http://www.sciencedirect.com/science/article/pii/S0957417411010761Expert Systems with Applications [Online], 2127-2131, 2012.

[4]B. Shim, K. Choi, Y. Suh, "CRM Strategies for a Small-Sized Online Shopping Mall Based on Association Rules and Sequential Pattern" Expert Syetem with Application, 7736-7742, 2012.

[5]Y. Zhang, "VIP Customer Segmentation Based on Data Mining in Mobile-Communications Industry" Computer Science and Education (*ICCSE*), International Conference ,156-159, 2010.

[6]G. M. Allenby, P. E. Rossi, "Marketing Models of Consumer Heterogeneity" *Journal of Econometrics*, 57-78,1998.

[7] G. Huaping, X. Qiong, "Study on Application of Customer Segmentation Based on Data Mining Technology" Future Computer and Communication, (*FCC*) , International Conference on. 167-170, 2009.

[8]J. Han, M.Kamber, J. Pei, "Data Mining Concepts and Techniques" Third Edition USA: Morgan Kaufman, *Elsevier*, 2012.

[9]S. H. Liao, Y. J. Chen, M. Y. Deng, "Mining Customer Knowledge for Tourism new Product Development and Customer Relationship Management" Expert Systems with Application, 4212-4223, 2010.