Ranking the Vehicle Insurance Customers by MADM Technique; a Case Study of Asia Insurance Company Branches in Zanjan Province

Mansour Momeni and Abdollah Nazari

Department of Management, Tehran University, Tehran, Iran.
Department of Management, Abhar branch, Islamic Azad University, Abhar, Iran.

ARTICLE INFO
Article history:
Received 3 April 2014
Received in revised form 25 July 2014
Accepted 20 September 2014
Available online 5 October 2014

Keywords:
Vehicle Insurance, Risk, Clustering, K-mean

ABSTRACT
In the race among the companies for winning the market, certainly getting or losing even one customer, is of special importance. The number of the companies and the race, makes it inevitable to make any possible effort to attract and get a given group of customers, as in this tough and intense competition, getting all the customers has become impossible. Hence, the companies feel that they need to segment and cluster the customers. What is discussed in the present paper is how to find and select the target customers and focusing on them, so that the companies could succeed. In our country, the vehicle insurance has the most important share in the insurance industry. In Iran, the auto body insurance price is determined on the basis of the Central Insurance tariff. The main purposes of the present study, are ranking the customers of the auto body insurance industry considering the risk level of each customer; helping the organizations and insurance companies in order to use the proper strategy for each and every group of customers; and improving the current position in the market. First and foremost the factors underlying the policyholders’ risk taking were detected by means of interviewing them and also by using the experts’ comments and reliable magazines’ surveys; and then the customers of vehicle insurance were clustered from two to ten clusters by means of K-mean method. By evaluating the validity index using the MADM technique, the best cluster number, was determined to be 5. Then, the clusters once again, were ranked by using the TOPSIS technique, considering the type of the indexes. The clustering method and ranking the customers, makes it possible for the managers to know their customer better and select the target customers and focusing on them, so that the decision-making process is more facilitated.

INTRODUCTION

One of the terms which is frequently used in the market clustering field, is market sector determination which involves determining the homogenous subsets of a mass market by using the customer clustering method on the basis of a set of variables. Therefore, the clustering is a sort of accumulation process in which the consumers with certain needs and expectations, are clustered in a certain sector of the market which is distinct from the other sectors. In clustering, it is assumed that the consumers which form a particular sector of the market, are homogeneous with each other, and inhomogeneous with other sectors. The market clustering shows a more rational and accurate compatibility of the product with the marketing efforts and consumers’ needs [1].

Nowadays, insurance industry is one of the most essential economic sectors of any country which is responsible for important tasks and functions. The insurance industry development in any country indicates the overall development and also shows the financial saving increases. In Iran, insurance is one of the most important factors of keeping and guaranteeing the capital [2]. Knowing the customers requires analyzing the target customers and clustering them which leads to detecting the groups of target customers; classifying them into main groups with similarities, and eventually putting these similar ones into one group [3]. The risk classification actually means clustering the customers with similar risk properties which have similar loss likelihood or impact occurrences. In most countries, the vehicle insurance prices are calculated on the basis of various demographic variables, the specifications of the vehicle and the policyholder’s loss record. Whereas in Iran, heedless to the aforementioned points that each of them involve several variables, the vehicle insurance price is determined on the basis of the Central Insurance tariffs. The result is that the customers with lower risk

Corresponding Author: Abdollah Nazari, Department of Management, Abhar branch, Islamic Azad University, Abhar, Iran.
Tel: 98 912 641 96 56; E-mail: Nazari_jea@yahoo.com
level can compensate the financial losses of the customers with higher risk; and therefore, there would not be significant difference among the customers with high and low risks, and the insurance organizations also are mistaken in detecting the customer groups and decision-making processes.

In order to determine the customers’ indexes and variables for clustering, firstly, the indexes were extracted from the reliable papers and books, and then considering the certain features of the vehicle insurance industry and using the viewpoints and comments of the experts and scholars of the field, the interfering variables were classified into four main categories of demographic features, vehicle specification, insurance policy features and the driver’s record. In the next step, a questionnaire was designed according to the detected factors by the experts of the field, who were selected of the many financial loss experts, bosses and managers of 10 insurance companies. Eventually, after analyzing the questionnaires, the final risk variables of the customers with certain scales were determined.

Clustering:
Classifying the similar things into several groups is one of the most important activities of human being. In the normal everyday life, this issue is a part of the learning process which as a result of the rapid development of the information technology and the increase of the stored data size in the databases, is getting more and more credit. Analysis of the stored data and converting them into usable information and knowledge for the organizations needs very powerful instruments. In marketing, the customers are firstly clustered according to the different indexes (variables). Then, the behavior of each sector is observed and planned for offering the more proper and more specialized service in each sector [4].

There are two main ways for clustering the customers:
1. Priority-based clustering, in which the researcher firstly selects the favorite cases in some of the variables, and then designs and classifies the customers according to the selected variables.
2. The clustering in which the researcher selects a set of interdependent variables, and then according to this selection, divides the customers into groups whose mean in case of the intergroup similarity is high and in case of intragroup similarity is in between [7].

The effectiveness of the clustering methods depends on the extent to which the selected sector is measurable, significant, accessible, observable, and excellent and applicable [8]. Finally the company, chooses one or several clusters and acts as if they are small markets [9].

In the present study, the customers of the auto body insurance, in an insurance company have been clustered according to their risk-taking level. The model suggested steps involve the variable definition, data gathering, data clearance, and model design and data analysis.

3.1. Variable Definition:
One of the important factors in making a model is the proper selection of the variables. In order to cluster the customers according to the risk-taking, the first and foremost step is to recognize the risk factors. The recognition and selection of the variables take place in two stages in the present study. In the first stage, the reliable scientific articles in journals from 2004 to 2011 were investigated. This stage is one of the most important stages of conducting this study. In this part, 16 variables were determined and extracted as shown in Table 1, and these variables were classified into four groups of demographic features, vehicle specifications, driver’s record, and insurance policy features in case of each type of the specifications. Then the features and specifications that cannot be accessed in all cases were omitted and finally 11 variables were determined for evaluating, shown in Table 1.

<table>
<thead>
<tr>
<th>Table 1: Found Factors from Experts’ Answers to Questionnaires and Interviews [12].</th>
</tr>
</thead>
<tbody>
<tr>
<td>Factors</td>
</tr>
<tr>
<td>----------------------------------</td>
</tr>
<tr>
<td>Demographic Features</td>
</tr>
<tr>
<td>Vehicle Specification</td>
</tr>
<tr>
<td>Driving Record</td>
</tr>
<tr>
<td>Insurance Policy Specification</td>
</tr>
</tbody>
</table>
In the second stage, a questionnaire was designed and answered by the experts. Considering this fact that the experts must be of the employees and managers of the insurance units loss field, and they must have enough knowledge and experience in the field; therefore the experts in the auto body insurance sector of the insurance companies and their managers were involved in the study. In our country, currently 25 insurance companies are active in vehicle insurance industry, of which all the experts, bosses and the managers of financial losses in ten insurance companies in Zanjan Province, were selected as the population of the study. According to the investigations done, the total number of the financial loss experts, bosses and managers of the units was estimated 432 people. Moreover, in order to determine the sample size, considering the fact that the population includes a finite but large number, and also considering this fact that the sampling was random, Cochran’s C test was used to calculate the sample size. By using the Cochran’s formula, the sample size was chosen 92 samples. In this regard, 92 questionnaires were distributed, proportionate to the market’s share of each of the ten insurance companies, and of these 92 distributed questionnaires, 81 were answered by the experts.

3.2. Data Collection:
In this study, the available data in the auto body insurance database of Asia Insurance Company were used and for the mentioned 11 characteristics, the data was extracted from the databases of Asia Insurance Company of Zanjan Province, for 120 people in a random style by using random numbers from 12876 automobiles, taxies and pickups.

The loss in auto body insurance generally fall into two separate groups of damages caused by accidents or the losses related to theft. The theft cases have led to this particular division of life places into two normal and crime hotspot areas in case of security. It should be noted that the auto body of the automobiles older than 15 years, cannot be legally insured, and also the insurance companies themselves are reluctant to insure the automobiles older than 10 years and taxies, and if they do so, the insurance prices of the mentioned cases would be higher.

3.3. Data Preprocess:
Data preprocessing is one of the important steps of clustering, and the way it is done significantly affects the accuracy of the whole process results. For preprocessing the data, the following actions is accomplished:
- Omitting the improper and lost data
- Converting all the qualitative amounts into quantitative ones
- Normalizing the amounts
- Extracting the Indexes’ Scales

### Table 2: Determining the Representative Clusters Using the Balanced Mean.

<table>
<thead>
<tr>
<th>Age</th>
<th>Gender</th>
<th>Location</th>
<th>License Type</th>
<th>License Record</th>
<th>Vehicle Use</th>
<th>Vehicle Color</th>
<th>Vehicle Age</th>
<th>Elapsed Distance (KM)</th>
<th>Damage Record</th>
<th>Insurance Payment</th>
</tr>
</thead>
<tbody>
<tr>
<td>C1</td>
<td>11.033</td>
<td>48.075</td>
<td>88.138</td>
<td>105.765</td>
<td>5.629</td>
<td>0.000</td>
<td>19.230</td>
<td>2.959</td>
<td>2.327</td>
<td>0.000</td>
</tr>
<tr>
<td>C2</td>
<td>12.523</td>
<td>57.690</td>
<td>0.000</td>
<td>115.380</td>
<td>6.050</td>
<td>89.747</td>
<td>89.37</td>
<td>8.937</td>
<td>6.153</td>
<td>410.257</td>
</tr>
<tr>
<td>C3</td>
<td>9.627</td>
<td>38.941</td>
<td>72.113</td>
<td>92.304</td>
<td>4.120</td>
<td>121.158</td>
<td>17.307</td>
<td>4.423</td>
<td>3.365</td>
<td>0.000</td>
</tr>
<tr>
<td>C4</td>
<td>10.495</td>
<td>53.711</td>
<td>79.572</td>
<td>115.380</td>
<td>4.712</td>
<td>46.421</td>
<td>26.524</td>
<td>6.856</td>
<td>5.102</td>
<td>222.813</td>
</tr>
<tr>
<td>C5</td>
<td>9.547</td>
<td>46.473</td>
<td>82.796</td>
<td>64.100</td>
<td>4.117</td>
<td>127.141</td>
<td>25.640</td>
<td>8.023</td>
<td>5.885</td>
<td>183.763</td>
</tr>
</tbody>
</table>

Clustering Analysis:
In this step, as the K-mean is used for clustering, and the clustering has been done in the best style and manner; hence, for each cluster, the mean of the records are calculated and for each and every cluster representative vector is drawn.

Now, in this stage, considering the eleven indexes, that in the first 7 indexes, have negative aspect in terms of customer risk, and in the next 4 indexes, the customer risk, brings about positive aspect, the Shannon entropy is used for calculating the weights of the indexes, and then via TOPSIS technique, 5 clusters are ranked as the representatives. The results is shown in Table 8. As it is obvious, the third cluster has the most risk and the worst status, and the first and fifth clusters are in the following positions, in this case.

### Table 3: Determining the Relative Closeness of Each Choice to the Ideal Solution.

<table>
<thead>
<tr>
<th>Cluster No.</th>
<th>CL*</th>
<th>Rank</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.3071</td>
<td>4</td>
</tr>
<tr>
<td>2</td>
<td>0.7690</td>
<td>1</td>
</tr>
<tr>
<td>3</td>
<td>0.1076</td>
<td>5</td>
</tr>
<tr>
<td>4</td>
<td>0.4652</td>
<td>2</td>
</tr>
<tr>
<td>5</td>
<td>0.3305</td>
<td>3</td>
</tr>
</tbody>
</table>
Conclusion:
In Iran, the auto body insurance price, is determined on the basis of the Central Insurance tariff, which leads to this particular result that the customers with lower risk level, can compensate the financial losses of the customers with higher risk; and therefore, there would not be significant difference among the customers with high and low risks. The purpose of this study is to cluster 120 customers of auto body insurance, on the basis of three groups of demographic index, driver’s record, and insurance policy features. The indexes selected were age, gender, driver’s residential location, driving license type, the use of vehicle, the color of the vehicle, the mileage per kilometer, the damage and loss record, and the insurance coverage level. By using the insurance experts’ comments and by means of paired comparison technique, ordinal and nominal indexes were converted into quantitative indexes, and they were made scale-less via Euclidian norm. Clustering insurance customers was done in several steps using Machaon Software, and after validation of the evaluating indexes thanks to the TOPSIS technique, the optimum number of clusters (5) was determined. Considering 11 indexes, where the first 7 indexes, have negative aspect in terms of customer risk, and the next 4 indexes, the customer risk, brings about positive aspect , the representative for each cluster was introduced by using the mean method , and then again the TOPSIS technique revealed the worst and the best cluster in terms of risk.

REFERENCES