Artículo de investigación

Using data mining techniques to extract profitability strategy

Usar técnicas de minería de datos para extraer la estrategia de rentabilidad

Usando técnicas de mineração de dados para extrair estratégia de lucratividade

Recibido: 9 de abril de 2018. Aceptado: 30 de mayo de 2018

Escrito por:

Zahedeh Jafari [4]
Zahra Jiriayi Sharahi[2]
Ali Dehbashi[3]

Abstract

Today, with the development of base systems and the high volume of data stored in them, there is a need for tools to process and use these data. The best tool is to discover rules between data, data mining and data mining techniques. In this article, we are looking for data mining tools to extract the profitability rules. In order to obtain the tangible results of this research, the information of customers of a restaurant in two parts has been collected through the distribution of a questionnaire. The first part is customer personal information, and the second part is the importance of each of the loyalty factors assigned by each customer. After collecting the data, the Excel file was imported and analyzed by Rapid Miner software, and clustering was performed on all customer information and customers were placed in clusters based on similarity of features and their behavior. Clustering is performed using the K-means algorithm. Then, an index is defined for categorization and is categorized on each cluster of clustering, and customers of each cluster are divided into two groups of profitable and non-profitable, and after extracting the rules from the decision tree, the strategies are presented for making more profit in the studied location.

Resumen

Hoy en día, con el desarrollo de sistemas básicos y el gran volumen de datos almacenados en ellos, existe la necesidad de herramientas para procesar y utilizar estos datos. La mejor herramienta es descubrir reglas entre datos, la minería de datos y las técnicas de minería de datos. En este artículo, buscamos herramientas de extracción de datos para extraer las reglas de rentabilidad. Con el fin de obtener los resultados tangibles de esta investigación, se obtuvo, en dos partes, la información de los clientes de un restaurante a través de un cuestionario. La primera parte es información personal del cliente, y la segunda parte es la importancia de cada uno de los factores de lealtad asignados por cada cliente. Después de recopilar los datos en un archivo de Excel, se exporta y analiza la información en el software Rapid Miner; se realiza la agrupación de toda la información del cliente en clústeres y se coloca en clústeres según la similitud de características y su comportamiento. La agrupación se realiza utilizando el algoritmo K-means. Luego, se define un índice para la categorización y se clasifica en cada grupo de clustering. Los clientes de cada clúster se dividen en dos grupos de rentable y no rentable, y después de extraer las reglas del árbol

[4] Ms. Industrial engineering, Pooyesh Qom higher education institute, Iran
email address: zahedeh.jafari@yahoo.com
[2] PhD. Candidate in industrial engineering, University of Yazd, Yazd, Iran
[3] Ms. MBA student, University of Semnan, Iran

Keywords: Customer relationship management, Clustering, K-Means, Classification, Decision Tree

de decisiones, se presentan estrategias para hacer más ganancias en la ubicación estudiada.

Palabras clave: gestión de relaciones con los clientes, agrupamiento, K-medias, clasificación, árbol de decisiones

Resumo

Atualmente, com o desenvolvimento de sistemas básicos e o grande volume de dados armazenados neles, há necessidade de ferramentas para processar e utilizar esses dados. A melhor ferramenta é descobrir regras entre dados, mineração de dados e técnicas de mineração de dados. Neste artigo, procuramos ferramentas de extração de dados para extrair as regras de lucratividade. A fim de obter os resultados tangíveis desta pesquisa, as informações dos clientes de um restaurante foram obtidas em duas partes através de um questionário. A primeira parte é a informação pessoal do cliente, e a segunda parte é a importância de cada um dos fatores de lealdade atribuídos por cada cliente. Depois de coletar os dados em um arquivo do Excel, as informações são exportadas e analisadas no software Rapid Miner; O agrupamento de toda as informações do cliente é feito em clusters e colocado em clusters de acordo com a similaridade de características e seu comportamento. O agrupamento é feito usando o algoritmo K-means. Em seguida, um índice é definido para categorização e classificado em cada grupo de agrupamento. Os clientes em cada cluster são divididos em dois grupos de lucrativos e não rentáveis e, depois de extraírem as regras da árvore de decisão, são apresentadas estratégias para obter mais lucros no local estudado.

Palavras-chave: gestão de relacionamento com clientes, agrupamento, K-médio, classificação, árvore de decisão

Introduction

Today, data are considered to be the heart of the business process of most companies, regardless of the wisdom and magnitude of the industry in all industries such as retailers, communications, manufacturing, facilities, transportation, insurance, credit cards and banking are aggregated through interactions in operational systems or for the discovery of knowledge. Therefore, there is a need for a tool that can process stored data and provide the user with information from this process.

Data mining can be a good tool for discovering this knowledge, and from there, in all businesses, profitability is the main purpose of its business executives. It can focus on data and use of Data mining tools to succeed in finding profitable business strategies by discovering hidden knowledge among the data.

Problem description

This research has been conducted to maintain the restaurant's customers with a profitability approach for the restaurant. Therefore, the interests and factors that lead to their loyalty should be considered, so that the restaurant can provide solutions to protect customers while maximizing profitability.

Goals and research assumptions

The following question are supposed to be answered by executing this research:

Question 1: Which data mining tools can simplify the search for a large number of customers (records)?

Question 2: Can the customer determine the profitability of the restaurant with the data mining tools?

Question 3: Can you extrapolate profitability strategies in this way?

Therefore, the main assumptions are:

- Customer clustering and determining the best K number in K-means from K = 4, 5, 6, 8
- Determine the most accurate categorization method among decision tree, KNN, perceptron methods.
- Classifying and determining customers profitably and non-profit.
- Determining the most accurate and profitable customer cluster
- Extracting tree rules and strategies.

Research history

After a lot of research on this topic with the sole purpose of the subject under consideration, we will refer to a summary of the articles in which the elements examined are common to the research.

First paper abstract: Customer Relationship Management is an increasing interest in a group of customers to change their behavior. Because the previous benefit from loyal customers is significant, so focusing on fugitive loyal customers can make a lot of profit. In this article, we focus on the loyal customers of a company and construct a model to predict partial escapes from loyal customers using three categorization models. Model Input (Statistical Society and Past Purchase Behavior). The benefits of this article are to re-engage loyal customers and prevent them from leaving, and turning partial escape into general outbreaks over time. This model can examine the possibility of escaping customers. This model can be used for all companies that operate in non-enclosed environments and it is difficult to detect customer escapes for them. This article uses categorization techniques in data mining. (Buckinx, W., & Poel, D. V. D. (2005).)

Second paper abstract: The core issue of customer relationship management is to prevent customers from falling through their maintenance. By minimizing customer depreciation, the company's profit will be maximized. This article proposes a hybrid design to deal with customer retention issues. It does this not only by predicting the risk of falling, but also by proposing conservation policies. The design involves two modes of use and training. In educational mode, it produces a loss-making model that can analyze and predict the likelihood of cultivators. It also creates a policy model that classifies variables by tagging each group based on their important features, and then classifying them in order to create the right policies for each customer's cluster is used. In use, the drop model is used to determine if the customer wants to leave the collection. If there is a high probability of leaving a customer, the policy model is used to suggest a specific customer-retention policy. The most important feature of this approach is that it not only anticipates the incineration but also exerts an active effort to reduce the incident. In addition to the predictor of this paper, this paper presents methods for reducing the crop. This article uses categorization techniques in data mining and focuses on the decision tree. (Keshtkar ,2017)

Third paper abstract: This paper proposes a new method based on which an RFM model adds a parameter, which calls it an extended RFM model, and joins the WRFM-based method with the K-means algorithm in data mining, with the hand Bringing the optimal K based on the Davis Boulder index is used and after classification, loyalty is created. The proposed method has shown that for clustering, when the developed model combines the WRFM with the K-means algorithm, it is possible to achieve a superlative improvement in classification accuracy in order to achieve an excellent CRM. The results of statistical tests for model validation show that the developed method for CRM has acceptable results with a high level of trust compared to other models that are commonly used by researchers. This paper focuses on clustering K-means techniques. (Seyed Hosseini ,M. , A. , Maleki, Gholamian ,M. ,2010)

Fourth paper abstract: This paper proposes an efficient CRM data framework and examines the effectiveness of two categorization models in data mining to predict customer behavior in the application of CRM. The use of standard marketing data sets and the use of the Bayesian classification is one of the benefits of

this article because of its ability to predict the likelihood that an instance belongs to a particular class. This method is used for large databases because of its high accuracy and stability to simple model learning. In order to satisfy the parameters (i.e., mathematical hope and variance) necessary for classification, classification requires only a small amount of educational data. It also categorizes real and discrete data. The focus of this paper is on the categorization techniques of the neural networks in particular. (Keshtkar and Talebi zadeh, 2017, Ming-Shian,T. , Keshtkar ,(2018))

Fifth paper abstract: Understanding and adapting to changes in customer needs requires survival in an environment that is constantly changing. In this paper, the use of data mining techniques, especially the decision tree, is suggested for data mining in the changes before and after a given point of customer needs. The proposed method for extracting variations is based on the decision tree, and can find interesting relationships between a large set of data that is useful for marketing campaigns. The proposed method automatically detects changes from customer profiles and sales data at different times. For each dataset, the decision tree is produced, the decision trees are compared with the matching rules, and it is difficult to answer the question as to what kind of change and how much it occurred. To overcome this problem, we set out rules for the decision tree, and we define three types of variations. Then, using matching rules, the similarities and differences are examined to determine the types of changes. The focus of this approach is on the customer retention dimension of customer relationship management, and uses categorization techniques. (Keshtkar and Dadkhoda zadeh, 2018)

Sixth paper abstract: This article proposes an energy forecasting system using fuzzy logic to reduce the failure, damage, and inefficiency of the factors of production. The proposed fuzzy logic approach helps the manufacturer predict the amount of energy consumption in the device based on inputs. These findings can provide solid foundations for the decision maker and analyst of the manufacturing system. Take appropriate strategies to ensure system efficiency and stability and make the right decisions using data management. The focus of this approach is on customer retention and uses categorization techniques (Leung, R. W. K., Lau, H. C. W., & Kwong, C. K. (2003).).

Research method

-The temporal and spatial realm of research. This research was conducted in a 6-month-long period. The research subject is a restaurant.

- Data collecting method and tool. Recognizing restaurant customers, determining profitable customers, and extracting strategies to earn more profit for the restaurant and retain customers. The goal of this research, which was carried out at the nonprofit Dynamics Institute. To collect data from real space, field search and questionnaire tool.

-Questionnaire design. Before distributing the questionnaire, the reliability and validity of the questionnaire were carried out and the results of calculations are given in clause 0. It should be noted that the questionnaire contains two tables, the first table is questioning the personal information of the customers, and the second table specifies the customer's view of the level of importance of the factors of loyalty to the restaurant.

-Validation of the questionnaire. Questionnaires were distributed among 31 respondents who were in the same position as the main respondents and Cronbach's alpha was calculated for the entire questionnaire. Whose value was calculated with 26 questions 0.411, which reached 0.653 in three steps.

1- Omitting the “praying room question” and reaching 0.563 Cronbach’s alpha value.

2- Omitting the “playing space for children”, placing it under “having special creativity” and reaching 0.616 Cronbach’s Alpha

value.

3- Omitting the "free internet", placing it under "having special creativity" and reaching .653 Cronbach's alpha value.

Then, the factor analysis was performed and it was determined that all questions of the questionnaire were divided into 8 groups. Therefore, the questionnaire has been grouped and Cronbach's alpha and factor analysis for these groups have been calculated separately. All groups have a minimum of alpha of 0.6 and have been set on a factor. The validity and reliability of the questionnaire were proved in this way. The calculation results are presented in Table 1. The tables of the final questionnaire were prepared after the changes were made to the format presented in Table 2 and Table 3 for distribution.

Table 1. Cronbach's Alpha Computation Results and Factor Analysis

#	Questions by group	Cronbach's Alpha	Factor Analysis
About food			
1	Food quality based on the standards of medical knowledge and health	0. 78	0. 877
2	Food price		0. 368
3	Having a visible kitchen		0. 812
Respecting customers			
4	Suitable food for sick people and individuals with a special diet	0. 715	0. 991
5	Elevator		0. 991
6	There is a responsible person in front of the customers at the catering lounge.		0. 271
Food serving based on principles			
7	Good cuisine for decorating and decorative items	0. 701	0. 827
8	Beauty decoration and convenient catering facilities		0. 865
9	Beautiful cuisine		0. 413
10	Diversity		0. 944
Restaurant hygiene principles			
11	Clean WC	0. 66	0. 844
12	Service discipline and the proper way of serving and serving the food		0. 844
Staff Responsibilities			
13	Observe the hygiene by the staff (gloves, etc.)	0. 713	0. 118
14	respecting the restaurant's apparent grooming (table and ...)		0. 998
15	dealing with and specializing personnel and respecting the principles and conventions of catering		0. 998
Vehicles			
16	Existence of proper parking space	0. 60	0. 818
17	Food delivery		0. 818
Physical properties			
18	Safe and safe staircase	0. 61	0. 813
19	Air conditioner		0. 813
Attractions			
20	Creating a certain superiority over other restaurants (creating entertainment to prepare a customer's order - Having a baby chair for children - Playing for children for children - Free Internet)	0. 768	0. 932
21	Enough or suitable for all tastes		0. 932
22	Relax the restaurant Or Relaxation Music		0. 786
23	Award And Discount		0. 56

Table 2. Questionnaire table 1

Table 1: Personal Customer Information
Please enter the option number in the last column, in the column labeled "Customer Response".

Criterion	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6	Customer's answer
Gender	Male	Female					
Age	Under 20	20-30	30-40	40-50	50-60	Over 60	
Education	Below diploma	diploma	Associate degree	B.S/B.A	MS. /MA.	PhD	
Income	Less than 800,000	800,000-1,000,000	1,000,000-1,500,000	1,500,000-2,000,000	2,000,000-2,500,000	more	
Marriage	Single	married					
Number of family members	1	2	3	4	5	More	
Days of visit	Thursday	Friday	Holidays	Working days	Without any rules	Others	
Special illness	Diabetes	Gastrointestinal Disorders	Blood Fat	Blood Pressure Oscillations	Other of more than one	No	
Visiting hour	12-14	14-16	16-18	18-20	20-22	22-24	
Accommodation in Qom	Qum Resident	Qum Suburbs	Passenger	Student in Qum			
Restaurant Position	Near home and far from workplace	Near workplace and far from home	Near Both	Far from both			
Diet Condition	Having special diet	Without special diet					
Employment Status	Government employee or retired	Private sector employee	Self employed	Student	Jobless, house wife, student	Having 2 jobs	
The number of ordered foods per visit (all round table guests)	1	2	3	4	5	More	
The number of visits per month	Less than 1	1	2	3	4	More	
The amount of the invoice	Less than 25,000	25,000-50,000	50,000-100,000	100,000-150,000	150,000-200,000	More	
How did you get to know the restaurant first?	Urban advertisement	Random visit	Acquaintance with personnel	Virtual space	Mouth to Mouth	Others	
Would you like to go to the restaurant to see if the restaurant is nearby or away?	Nera	Far	Not important	I prefer to eat restaurant food at home			

Table 3. Questionnaire table 2

Table 2: Loyalty Factors and Their Significance Please indicate the level of importance of each level of loyalty from your point of view.							
#	Loyalty factor	1 Very important	2 Important	3 Average importance	4 Low importance	5 Very low importance	**Customer answer**
1	Quality of food based on the standards of medical knowledge and health						
2	Special treatment of personnel and observance of principles and conventions of food						
3	Food prices						
4	Suitability of food for sick people and people with special diet						
5	Have a suitable parking or special parking place						
6	Adhere to the restaurant's cleanliness (table, etc.)						
7	Beauty decoration and convenient dining utensils						
8	Cedar suitable for food and Decorating and Doricine						
9	Observance of sanitation by the staff						
10	Be responsible for responding to customers at the catering lounge						
11	Having a kitchen visible to customers						
12	Air conditioning						
13	Relax the restaurant or play relaxing music						
14	Send food with a courier to the home						
15	Creating a certain superiority over other restaurants (creating entertainment to prepare a customer's order - having a baby chair for children - playing space for children - free internet)						
16	lifts						
17	Safe and safe staircase						
18	Light enough or suitable for all tastes						
19	Restaurant Diversity						
20	Service discipline and the proper way of serving and serving the food						
21	Prizes and discounts						
22	Clean kitchens						
23	Beautiful cuisine for meals						

- The statistical community has gathered the temporal and spatial realm.500 clients of the restaurant were surveyed during the one-month period from July 8 to August 8 in all the working hours of the studied restaurant (named in the abstract). After confirmation of the results of the validation of the questionnaire, the questionnaires were provided to the restaurant registration personnel and provided to them by customers. Every time a person registers his order, he will have time to complete the questionnaire and deliver it to the restaurant staff until the moment he arrives to pay the bill to the fund. 500 questionnaires have been submitted to the restaurant for a period of one month.

- Analysis method.To analyze the data and discover the knowledge and hidden pattern among them, the Quick Miner 5 data mining software is used as described.

- Entering data in table in Excel file.After receiving completed questionnaires, a table has been prepared in Excel software, its rows for each customer and its columns are 18 questions in Table 2 and 23

in Table 3. Therefore, we have 500 records. The data collectively will constitute a table of 500 * 42.

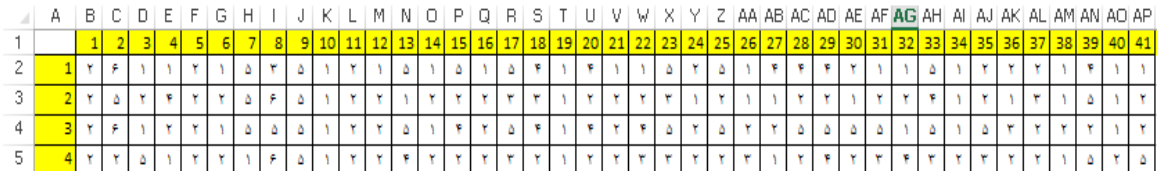

Fig.1. An overview of how to enter data in the table

- Performing clustering operation on the data. Clustering operations are performed on data in Figure 1, five hundred records. The purpose of the clustering is to place customers with similar features and behaviors in a group and to simplify the examination of a large number of customers. In other words, we are looking for customers that have the same behavior based on the factors studied, along with each other in a cluster. The K algorithm is the mean of the algorithm for clustering. Which is done using the quick miner 5 software on the data by the operator K-means from the subcategory of clustering operators. The clustering model in the software is presented in Fig. 2.

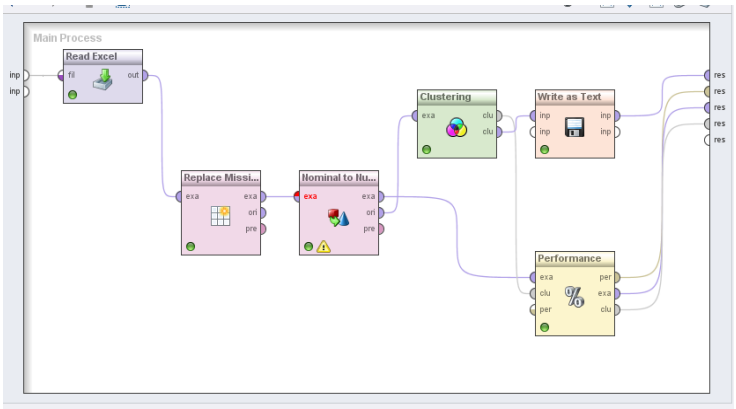

Fig. 2. Clustering process in Rapid Miner 5 software

The clustering operation for these five hundred records has been performed 4 times, with the K number being implemented in Algorithms 4, 5, 6, 8. The number of records in each cluster is divided by the value of K in Table 4.

Table 4. Number of records in each cluster by K value

Cluster Number	Cluster 1	Cluster 2	Cluster 3	Cluster 4	Cluster 5	Cluster 6	Cluster 7	Cluster 8	Davis Index	K value in clustering
Number of members in each	159	192	7	142					-2.490	**4**
	84	108	7	186	115				-2.417	**5**
	82	67	149	110	7	85			-2.467	**6**
	39	87	7	55	101	83	52	76	-2.512	**8**

The goal of clustering is to simplify the behavior of five hundred customers.

- *Data preparation for clustering operation.* Now, each of the four above modes is done in the manner described below, and the results are compared together. First, we define the indicator for profit (Aleksandrovna Maximova and Aleksandrovich Belyaev, 2017). This index is:

Profitability index = Number of visits per month * (Number of people around the table/ Invoice amount) (1)

Thus, the index mentioned in relation (1) to all five hundred data entered in accordance with Figure 1, and to records whose indexes are less than the average index column, the 0 is labeled and the records whose indexes are above the average is the 1 label belonging Takes place. Then the records of each cluster are separated, so the data will be ready to be categorized.

- *Comparison between different clustering methods*. To select the best classification technique on each cluster, KNN, perceptron, decision tree methods are selected and the precision of each method is calculated using the fast miner 5 software, and the results are shown in Tables 5, 6, 7 and 8.

Table 5. Different clustering methods' accuracy for K=4

#	Cluster Number	Decision Tree Accuracy	KNN Accuracy	Perceptron Accuracy
1	0	91.83	75.50	88.08
2	1	82.89	70.79	85.92
3	2	Missing data cluster has been deleted.		
4	3	91.48	76	78.05
Average		88.733	74.09	84.01

Table 6. Different clustering methods' accuracy for K=5

#	Cluster Number	Decision Tree Accuracy	KNN Accuracy	Perceptron Accuracy
1	0	83.33	86.11	85.97
2	1	88.91	72.27	85
3	2	Missing data cluster has been deleted.		
4	3	82.72	73.68	83.86
5	4	93.94	77.50	79.09
Average		87.225	77.39	83.48

Table7. Different clustering methods' accuracy for K=6

#	Cluster Number	Decision Tree Accuracy	KNN Accuracy	Perceptron Accuracy
1	0	88.89	80.56	85.42
2	1	77.86	88.10	82.62
3	2	89.95	74.48	81.19
4	3	79.09	76.36	82.73
5	4	Missing data cluster has been deleted.		
6	5	91.67	74.17	82.22
	Average	85.49	78.734	82.836

Table 8. Different clustering methods' accuracy for K=8

#	Cluster Number	Decision Tree Accuracy	KNN Accuracy	Perceptron Accuracy
1	0	90	82.5	82.5
2	1	87.36	71.39	80.42
3	2	Missing data cluster has been deleted.		
4	3	81.33	77.67	83.67
5	4	95	73.09	85
6	5	83.61	74.58	79.72
7	6	84.67	78.67	82.67
8	7	92.14	63.04	60.18
Average		87.73	74.42	79.16

- Comparison of average accuracy of each clustering method and selecting the better method**.** The most precise method for categorizing is based on the results stated in Table 9 of the Decision tree method. Which is optimal according to the Davis index of K=5 clusters.

Table 9. Average accuracy of different clustering methods for different K values

K	Decision tree	KNN	Perceptron
K=4	88.733	74.09	84.01
K=5	87.225	77.39	83.48
K=6	85.49	78.734	82.836
K=8	87.73	74.42	79.16
Average	87.29	76.1585	82.3715

- ***Performing clustering*****.** At this stage, the classification was performed on all five cluster clusters K = 5 through the decision tree, and the prediction of the software was determined by the process of Figure 3 for the label of each record in each cluster.

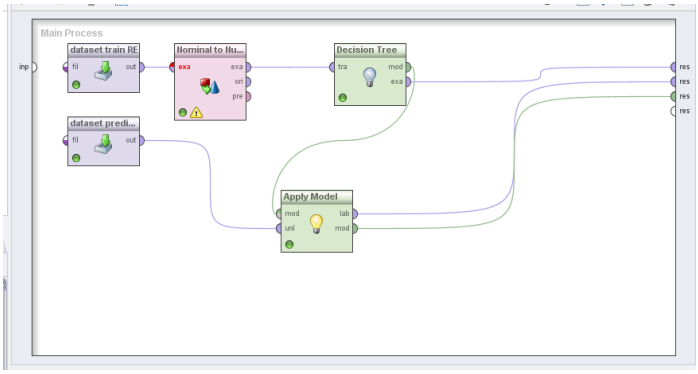

Fig.3. Clustering process

- ***Calculation of profitable and non-profitable customers' percent for each cluster*****.** In clustering with K = 5, percent of the profitable customers per cluster are calculated and in descending order the percentage of profitable customers is as follows: Clusters 2, Clusters 4, Clusters 5, Clusters 1. After

strategy discovery, rules of the cluster 2 can generate more revenue for restaurant. The rules of cluster 1 are described in Table 10 which is derived from decision tree 4 according to the process.

A4 :Revenue
A8 :Special illness
A14 :Number of ordered foods
A15 :Number of visits per month
S4 :Food suitability for sick people
S16 :elevator

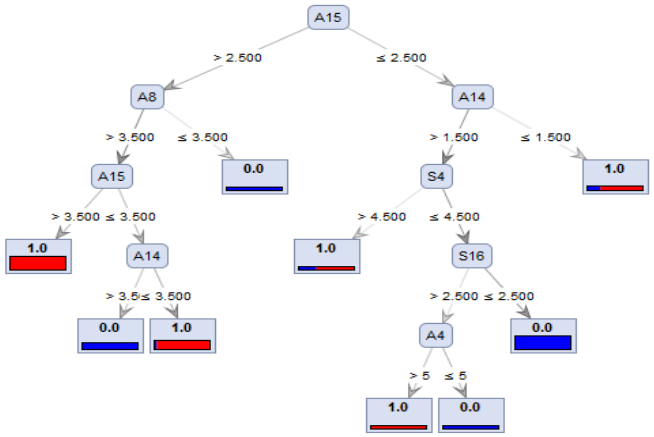

Fig. 4. Decision tree 2 of the clustering (K=5)

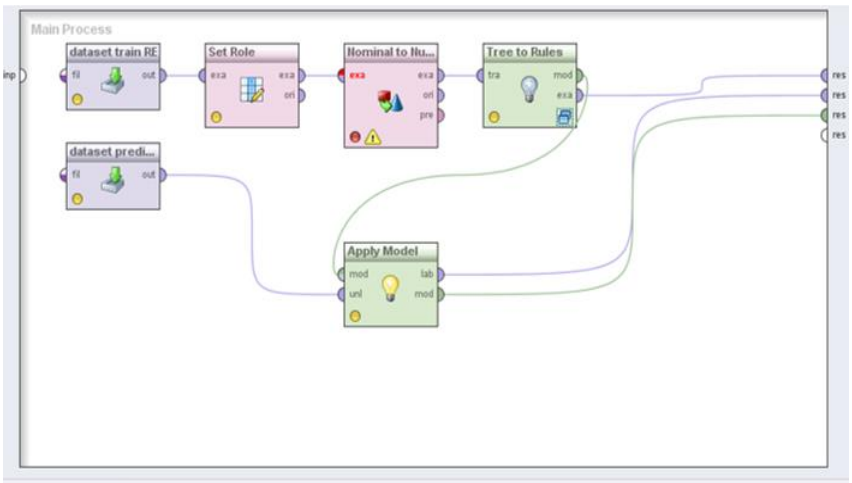

Fig.5. Rules extraction process in Rapid Miner Software

Table 10. Rules of cluster 2 from clustering with K=5

#	Rules	Profitability or non-profitability
1	if visits per month> 2.500 and special illness> 3.500 and visits per month > 3.500 then 1.0 (0 / 33)	Profitable
2	if visits per month > 2.500 and Special illness> 3.500 and visits per month ≤ 3.500 and number of ordered foods per visit> 3.500 then 0.0 (10 / 0)	Non-profitable
3	if visits per month > 2.500 and Special illness> 3.500 and visits per month ≤ 3.500 and number of ordered foods per visit≤ 3.500 then 1.0 (1 / 16)	Profitable
4	if visits per month > 2.500 and Special illness≤ 3.500 then 0.0 (3 / 0)	Non-profitable
5	if visits per month ≤ 2.500 and number of ordered foods per visit> 1.500 and food being suitable for sick people and people with special illness > 4.500 then 1.0 (1 / 2)	Profitable
6	if visits per month ≤ 2.500 and number of ordered foods per visit> 1.500 and food being suitable for sick people and people with special illness ≤ 4.500 and elevator> 2.500 and Salary> 5 then 1.0 (0 / 2)	Profitable
7	if visits per month ≤ 2.500 and number of ordered foods per visit> 1.500 and food being suitable for sick people and people with special illness ≤ 4.500 and elevator> 2.500 and Salary≤ 5 then 0.0 (3 / 0)	Non-profitable
8	if visits per month ≤ 2.500 and number of ordered foods per visit> 1.500 and food being suitable for sick people and people with special illness ≤ 4.500 and elevator≤ 2.500 then 0.0 (33 / 0)	Non-profitable
9	if visits per month ≤ 2.500 and number of ordered foods per visit≤ 1.500 then 1.0 (1 / 3)	Profitable

Correct: 105 out of 108 training examples. Coverage= 0.972, confidence= 88.91

Results

- Extracting the Priority of Loyalty Factors from the Perspective of Lucrative Cluster Customers. The priority was the most important factors of loyalty from their point of view for the investment of the owner of the restaurant according to Table 11:

Table 11. Prioritizing the most important factors from the perspective of profitable customers

#	Profitable group interest	The percentage of very important votes of customers for each factor in cluster 2	Priority
1	Quality of food based on the standards of medical knowledge and health	45/59=0.762	4
2	Special treatment of staff and observance of principles and conventions of food	42/59=0.711	6
3	Food prices	23/59=0.389	15
4	Suitable food for patients and people with special diet	17/59=0.288	16
5	Existence of a suitable parking place or special parking	27/59=0.457	13
6	observing the restaurant's apparent grooming (table and ...)	52/59=0.881	1
7	beauty decoration and convenient dining	40/59=0.677	7
8	cuisine suitable for food and decorations	37/59=0.627	9
9	Respect for health personnel by staff	49/59=0.830	2
10	There is a responsible person for our clients in the catering lounge	40/59=0.677	7
11	Having a kitchen visible to customers	24/59=0.406	14
12	Air conditioning	39/59=0.661	8
13	Relaxing restaurant atmosphere or playing relaxing music	33/59=0.559	11
14	Delivering th food to the house by courier	23/59=0.389	15

15	Creating a certain superiority over other restaurants (creating entertainment to prepare a customer's order - Having a baby chair for children - Playing space for children - Free Internet)	24/59=0.406	14
16	elevator	33/59=0.559	11
17	Convenient and safe staircase	30/59=0.508	12
18	Suitable light for all tastes	39/59=0.661	8
19	Restaurant Diversity	43/59=0.728	5
20	Service provision discipline, proper service method and hygienic food serving	48/59=0.813	3
21	Awards & Discounts	16/59=0.271	17
22	Clean WCs	48/59=0.813	3
23	Beautiful outdoor space for Meals	36/59=0.610	10

Considering the most important factors in the order of priority can help the manager of the restaurant to make profit.

-Extraction of the number of profitable clients divided by the options of the first table of the questionnaire in the cluster 2. After identification of the interests and tastes of profitable customers, the restaurant manager (study unit) can, by implementing their interests, in various ways, in addition to preserving them, to help their profitability, so it is likely to come. In the implementation of the interests, there is a need for customer attributes (Brunner and Ganga-Contreras, 2017). For example, the age group's focus on implementing the loyalty factor of priority 13 can be effective. When we know that a large number of profitable clients in the age group 20 to 40 years old. Therefore, the interests of this age group can be noticed.

Table 12. Feature of Profitable Customers Separated in Cluster 1

#	Customer's Feature	Option 1	Option 2	Option 3	Option 4	Option 5	Option 6
1	Sex	43	15	0	1	0	**0**
2	Age	2	21	20	10	3	**3**
3	Education	4	7	7	22	12	**7**
4	Salary	0	3	9	12	10	**25**
5	Marital Status	8	48	0	0	0	**3**
6	Number of people	5	22	17	6	7	**2**
7	How to visit	0	3	9	3	36	**8**
8	Special illness	0	0	0	0	1	**58**
9	Visit hours	9	6	0	0	24	**20**
10	Accommodation in Qom	48	0	10	1	0	**0**
11	Location of the restaurant	12	11	7	26	3	**0**
12	Diet Status	2	56	1	0	0	**0**
13	Employment status	10	16	20	4	7	**2**
14	Number of ordered foods	13	25	9	4	2	**6**
15	Number of visits per month	4	5	17	11	7	**15**
16	Invoice amount	4	21	20	5	5	**6**
17	How to know the restaurant	4	6	4	0	0	**12**
18	Preferred remoteness	11	6	35	4	3	**0**

Suggestions for the future

It is suggested that in order to get the most out of this article, the contents of the questionnaire were designed for each production unit designed and fitted to the same location, validated by the questionnaire, and the data was collected and used in accordance with the method used. In addition, in order to get better results, we can compare the results of several of the following methods, and each one that leads to the extraction of more logical rules and rules for a restaurant or business unit:

- Initially, clustering can only be done on loyalty factors, and then the data corresponding to the personal information of the customers are extracted for each cluster's IDs, then the grouping operations are performed and the results are obtained. Therefore, you can compare the two methods using the coverage percentage of the rules and the percentage of confidence.
- It is also suggested that in order to select the best number of clustering, ie, the value of K, it should be considered that the number of clusters is not excessive enough to make investigations difficult. Because the goal of clustering is to simplify the investigation of a large number of customers, the higher their number and the time interval of data collection, the better and more realistic results are obtained. Therefore, we can use other indicators to determine optimal clustering.

References

Aleksandrovna Maximova, O & V. Aleksandrovich Belyaev (2017). Generational Indigenation in a Multi-Ethnic and -Religious Society (Tatarstan, Russia). Opción, Año 33, No. 84 (2017): 38-64.

Bong-Horng, C., Ming-Shian,T. , Cheng-Seen ,H. ,(2007) ,Toward a hybrid data mining model for customer retention, Knowledge-Based Systems 20, 703–718.

Brunner, J.J & Ganga-Contreras, F. (2017). Vulnerabilidad educacional en América Latina: Una aproximación desde la sociología de la educación con foco en la educación temprana Opción, Año 33, No. 84 (2017): 12-37.

Buckinx, W., & Poel, D. V. D. (2005). Customer base analysis: Partial defection of behaviorally-loyal clients in a non-contractual FMCG retail setting. *European Journal of Operational Research*, 164, 252–268.

G Soleimani, M Amiri, SM Khatami, MJ Isfahani :Using S Technology, in the Automotive Industry, with the Approach of Its Implementation in Commercial Vehicles . Industrial Engineering & Management Systems.2016; 15(4): pp.290-297.

Keshtkar (2018). Numerical analysis of transcritical carbon dioxide compression cycle: a case study, 7 (1), 1-6.

Keshtkar M. M., (2017). Energy, exergy analysis and optimization by a genetic algorithm of a system based on a solar absorption chiller with a cylindrical PCM and nano-fluid, Int. Journal of Heat and Technology, 35 (2), 416-420.

Keshtkar M. M., Talebizadeh, P., (2018). Investigation of transient conduction–radiation heat transfer in a square cavity using combination of LBM and FVM, Sadhana, 43 (4), 145-155.

Keshtkar M. M.; Dadkhoda Zadeh M. (2018). Thermal Simulation of the Symmetric and Asymmetric Arrangement of Barriers on Heat Transfer Enhancement in a Porous Gas Heat Exchanger, Journal of Thermal Science and Engineering Applications, 10 (1), 120-135.

Kim, J. K., Song, H. S., Kim, T. S., & Kim, H. K. (2005). Detecting the change of customer behavior based on decision tree analysis. Expert Systems with Applications, 22, 193–205.

Leung, R. W. K., Lau, H. C. W., & Kwong, C. K. (2003). On a responsive replenishment system: A fuzzy logic approach. Expert Systems with Applications, 20, 20–32.

Seyed Hosseini ,M. , A. , Maleki, Gholamian, M. , Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty, (2010),Expert Systems with Applications 37, 5259–5264Bahari , F . , Elayidom , S,. (2015). An Efficient CRM-Data Mining Framework for the Prediction of Customer Behavior, Procedia Computer Science, 46, 725 – 731.