International Journal of Industrial Engineering & Production Research March 2022 Vol. 33, No. 1: 1-16
DOI: 10.22068/ijiepr.33.1.1

RESEARCH PAPER

Customer Clustering Based on Factors of Customer Lifetime Value with Data Mining Technique (Case Study: Software Industry)

Elaheh Bakhshizadeh[1], Hossein Aliasghari[2], Rassoul Noorossana[3] & Rouzbeh Ghousi[4]

Received 8 October 2020; Revised 13 November 2021; Accepted 30 November 2021;

ABSTRACT
Organizations have used Customer Lifetime Value (CLV) as an appropriate pattern to classify their customers. Data mining techniques have enabled organizations to analyze their customers' behaviors more quantitatively. This research has been carried out to cluster customers based on factors of CLV model including length, recency, frequency, and monetary (LRFM) through data mining. Based on LRFM, transaction data of 1865 customers in a software company has been analyzed through Crisp-DM method and the research roadmap. Four CLV factors have been developed based on feature selection algorithm. They also have been prepared for clustering using quintile method. To determine the optimum number of clusters, silhouette and SSE indexes have been evaluated. Additionally, k-means algorithm has been applied to cluster the customers. Then, CLV amounts have been evaluated and the clusters have been ranked. The results show that customers have been clustered in 4 groups namely high value loyal customers, uncertain lost customers, uncertain new customers, and high consumption cost customers. The first cluster customers with the highest number and the highest CLV are the most valuable customers and the fourth, third, and second cluster customers are in the second, third, and fourth positions respectively. The attributes of customers in each cluster have been analyzed and the marketing strategies have been proposed for each group.

KEYWORDS: *Customer lifetime value, LRFM model, Data mining, Clustering.*

1. Introduction

Competition and complicity in businesses have made organizations develop their innovative measures to meet the market needs and increase customer satisfaction [1]. Mass marketing approach is not able to meet diverse customer needs. Therefore, this diversity must be met by clustering customers who have the same needs and purchase behaviors. A suitable clustering helps organizations identify their customers well and deliver them proper goods and services appropriately; this, as a result, will guarantee good relations with their customers [2]. In fact, clustering is a technique to identify value-creating and profitable customers [3]. Organizations recognize their customers differently based on their values. Value-based clustering differentiates between customers according to the profit that they bring and the cost of establishing and preserving relations with them [4].

Customer lifetime value has been paid attention to as a valuable concept in general marketing and management to increase benefit. In other words, customer lifetime value is the current value of the customers' incoming profit [5]. This concept is a decisive criterion to improve customer relationship management; besides, adjusting organizations' policies based on customer lifetime value is the ultimate goal of marketing, according to a large numbers of experts [6]. Moreover, organizations adopt customer lifetime value to cluster customers, determine the target market, allocate resources, analyze profitability, formulate strategies, manage risk, and etc. [7].

Organizations analyze their customers' data to determine their preferences. As a result, managers can be able to enhance their decision

* Corresponding author: *Rassoul Noorossana*
rassoul@iust.ac.ir

1. *Departement of Business Management and MBA, Farabi Campus, University of Tehran.*
2. *Industrial Engineering Department, Iran University of Science and Technology*
3. *Industrial Engineering Department, Iran University of Science and Technology.*
4. *Industrial Engineering Department, Iran University of Science and Technology.*

support systems. In this regard, data mining is a practical, well-known and trustworthy method to analyze data and extract knowledge that helps marketing decisions. Considering marketing approach, different customer lifetime value and various purchase behaviors of the customers make it illogical to adopt the same strategy for all of them. Therefore, it seems necessary to cluster customers through data mining and devise appropriate strategies accordingly [1].

The present case study, a software company, is not an exception in this regard and has to cluster its customers to make proper decisions about customer relationship management and avoid overspending on marketing. This study aims to cluster the customers of the software company based on customer lifetime value (LRFM Model) with data mining technique. Using CLV formula in the next step, customer lifetime value of each cluster will be calculated and then clusters and their factors in each cluster will be ranked. Finally, all the clusters will be analyzed and the appropriate strategies will be formulated accordingly.

The literature on customer segmentation, customer lifetime value, related models (particularly LRFM model), and data mining and clustering techniques have been considered in the beginning of the present research. What follows next covers the methodology that is based on Crisp-DM. Finally, the research roadmap will be presented.

2. Literature Overview

In this section, firstly, the concept of customer segmentation and the methods through which customers could be segmented are described. Then, Customer Lifetime Value (CLV), its models, and particularly LRFM model as a segmentation method are explained. It is followed by a brief description of data mining and clustering techniques. This section finishes with the related works which investigate the previous research on this area.

2.1. Customer segmentation

Fu et al. [8] believed that segmentation is the process of classifying the heterogeneous macro-market into homogenous micro-market according to the same needs. This process helps the organizations differentiate between various customer groups and apply suitable strategies toward satisfying customers' needs. The results of such a process can be used in different marketing fields such as product, design and development, brand management, product distribution, pricing, resource allocation, and devising marketing plans [9].

Customer segmentation brings two considerable benefits to the management and marketing departments; the first benefit is all about identifying the main customer groups with the highest level of loyalty and profitability. In this regard, decision-makers are able to target high value customers. The second benefit enables the organizations to gain knowledge of customers and their purchase behaviors so that each customer group will be delivered appropriate services based on their needs and preferences. In addition, the knowledge that is gained from customer segmentation paves the way for a better resource allocation to customer groups [10].

In this regard, geographic, demographic, psychological, psychographic or behavioral criteria have been presented in order to segment customers [11]. Value-based segmentation, which is currently of great importance, uses factors of customer lifetime value to identify profitable customers and formulate appropriate strategies for each target group [12].

2.2. Customer lifetime value

Das et al. [13] believed that customer lifetime value is one of the most common and accepted model of customer evaluation that includes present value of the cash flows generated by customers. Today, CLV forms the basis of marketing data evaluation with the goal of presenting marketing patterns and standards. Providing knowledgeable experts with big datasets may enable the organizations to classify the customers in order of their profitability .This classification will guarantee long-lasting and profitable relationships with customers [4].

Estrella-Ramon et al. [14] stated that value-based classification not only meets the customers' needs, but also satisfies the organization's wants since the organization may be able to allocate its resources more accurately to the customers, resulting in the improved financial performance.

They also believed that customers were classified according to three criteria: 1. classification based on evaluating the amount of CLV. 2. Classification based on the conceptual components of CLV such as current value, potential value, and customer loyalty. 3. Classification based on CLV with use of purchase information, customer transactions and other supplementary information.

2.3. CLV models

Gupta et al. [15] presented various CLV models as follows: RFM models, probability models, econometric models, persistence models, computer science models, and diffusion/growth models.

Hughes [16] presented RFM model for the first time to classify and recognize the customer value. This model consists three dimensions namely recency, frequency, and monetary and is used to identify customer behaviors and analyze their future purchases. Recency is the time interval since the last purchase (based on day or month). If the customers have purchased something recently, they are more inclined to do so again. Frequency, as an important indicator of customer loyalty, is the number of purchases in a period of time. In other words, the more the customers buy, the more they become loyal. Monetary is the amount of money that's paid in a period of time for each visit, showing the level of customer participation; the more the customers pay, the more they participate in increasing company income [10] , [17].

RFM model has been developing since two decades ago. Accordingly, Chang and Tsay [18] developed this model to LRFM by adding length as the fourth dimension. Length is about the period of time between the first and last contract of the customers with the company. Needless to say, length plays a crucial role in customer loyalty.

Considering all these four dimensions, we can use a matrix with 5 groups and 16 sub-groups to cluster the customers. Such a matrix is shown in Fig. 1 and includes:

1) Core customers including:
- high value loyal customers: LRFM $\boxed{\uparrow\downarrow\uparrow\uparrow}$
- high frequency buying customers: LRFM $\boxed{\uparrow\downarrow\uparrow\downarrow}$
- platinum customers: LRFM $\boxed{\uparrow\downarrow\downarrow\uparrow}$

2) Potential customers including:
- potential loyal customers: LRFM $\boxed{\uparrow\uparrow\uparrow\uparrow}$
- potential high frequency customers: LRFM $\boxed{\uparrow\uparrow\uparrow\downarrow}$
- potential consumption customers: LRFM $\boxed{\uparrow\uparrow\downarrow\uparrow}$

3) Lost customers including:
- high value lost customers: LRFM $\boxed{\downarrow\uparrow\uparrow\uparrow}$
- frequency lost customers: LRFM $\boxed{\downarrow\uparrow\uparrow\downarrow}$
- consumption lost customers: LRFM $\boxed{\downarrow\uparrow\downarrow\uparrow}$
- uncertain lost customers: LRFM $\boxed{\downarrow\uparrow\downarrow\downarrow}$

4) New customers including:
- high value new customers: LRFM $\boxed{\downarrow\downarrow\uparrow\uparrow}$
- frequency promotion customers: LRFM $\boxed{\downarrow\downarrow\uparrow\downarrow}$
- spender promotion customers: LRFM $\boxed{\downarrow\downarrow\downarrow\uparrow}$
- uncertain new customers: LRFM $\boxed{\downarrow\downarrow\downarrow\downarrow}$

5) Consuming resource customers including:
- low consumption cost customers: LRFM $\boxed{\uparrow\downarrow\downarrow\downarrow}$
- high consumption cost customers: LRFM $\boxed{\uparrow\uparrow\downarrow\downarrow}$

Later, the average LRFM values of each cluster must be compared with total average LRFM values of all clusters. If average LRFM values of each cluster are greater than total average LRFM values of all clusters, the arrow will be upright and vice versa [19].

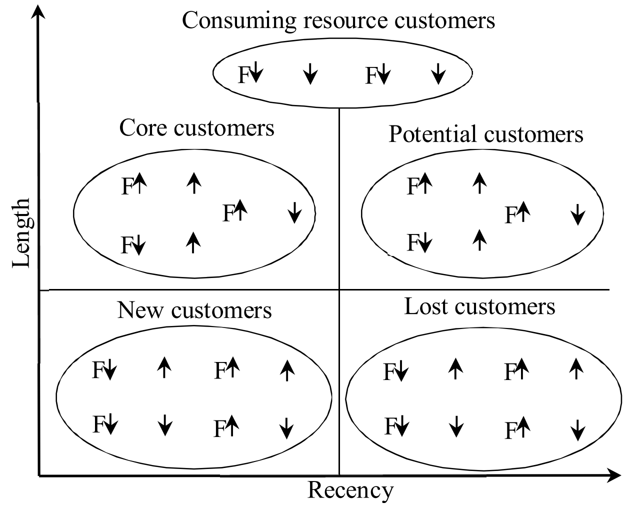

Fig. 1. Customer clustering based on LRFM

2.4. Data mining

To be compatible with the changing world, most of the organizations attempted to benefit from technological advances. Despite the technological advances, those organizations were not able to acquire knowledge from their data bases. Therefore, data mining was used as a powerful tool to extract knowledge and patterns from data [20]. Chen et al. [21] stated that data mining is the process of extracting hidden, unknown, and useful knowledge from the data. Cakir et al. [22] defined data mining as a process consisting of selecting, discovering, and modeling of huge amounts of data in order to discover hidden patterns and relations. Furthermore, data mining is a hybrid technique that integrates databases, statistics, machine learning, signal processing, and high-performance computing. It also fulfills various functions such as prediction, clustering, classification, and association [23].

Clustering, as one of the most important data mining methods, is a process that classifies data into meaningful groups namely clusters. Data are clustered through maximizing similarity within groups and minimizing similarity between groups [10]. This method is highly practical due to its ability to manage huge amounts of data. K-means is one of the most common clustering algorithms and considered as an unsupervised machine learning technique. Based on this algorithm, n elements will be divided into k clusters with the maximum intragroup similarity and intergroup difference [11].

2.5. Related works

During recent years, a large number of researches have been carried out to investigate CLV-based clustering or segmentation of customers through data mining algorithms. Ait daoud et al. [24] conducted a research to present a model capable of investigating and predicating customer churn in e-business. To do so, the customers were classified into 7 clusters using LRFM model and k-means algorithm. Using simple decision tree, artificial neural networks and decision tree ensemble, the customers were later divided into three groups namely fully reluctant, relatively reluctant, and loyal. Chiang [25] investigated customer value in airlines. Inspired by RFM model, a new model called FSLC was then presented consisting of frequency, seasons, locations of travelling, and cancellation times. In the next step, the passengers were segmented through ward's method. Dursun and Caber [26] conducted a research on profiling profitable hotel customers. To identify and segment the customers, RFM model, k-means, and SOM algorithms were used. Wei et al. [3] segmented the children in dentistry clinics through LRFM. Using SOM in the next step, the patients were classified into 12 main clusters. Li et al. [19] conducted a cluster analysis on textile industry that included two steps: ward's method for determining the number of cluster and k-means for clustering. Based on RFM model, Liang [27] developed a model to analyze customer value in car maintenance industry. Clustering results, which were obtained by using k-means and SOM, showed that the customers were classified into three main groups. Cheng and Chen [28] carried out a research on customers in electronic industry through RFM model. In this research, k-means algorithm was used to cluster the customers; by considering Rough Set Theory and LEM (learning from examples module) algorithm, segmentation rules were evaluated to help the organization improve its CRM.

3. Methodology

3.1. Crisp-DM

Crisp-DM is a data mining method that has been used in this study. Crisp-DM is a complete data mining model that presents methods of identifying the input and output of any process [29]. As shown in Fig. 2, this method consists of six phases namely business understanding, data understanding, data preparation, modeling, deployment, and evaluation [30].

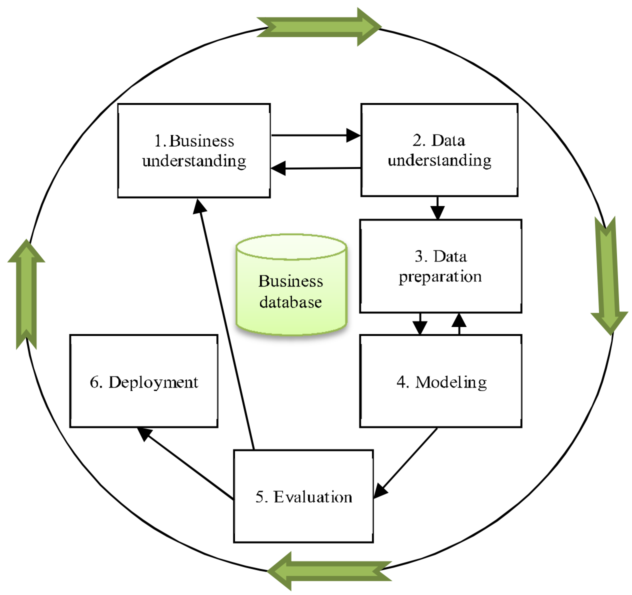

Fig. 2. Crisp-DM

Business understanding phase includes understanding business objectives, situation assessment, and determining data mining goals. Data understanding phase starts with data collection and continues with operations that lead to data identification. Data preparation phase covers all the necessary steps to build the ultimate data sets gained from primary data. In the modeling phase, modeling techniques are chosen and their components will be calibrated. Evaluation phase deals with the analysis of the model. The final phase includes applying the appropriate model to the decision-making processes and presenting strategies [31] , [32].

3.2. Research roadmap

The research roadmap is designed in a way that all the research steps are in accordance with Crisp-DM (Fig. 3).

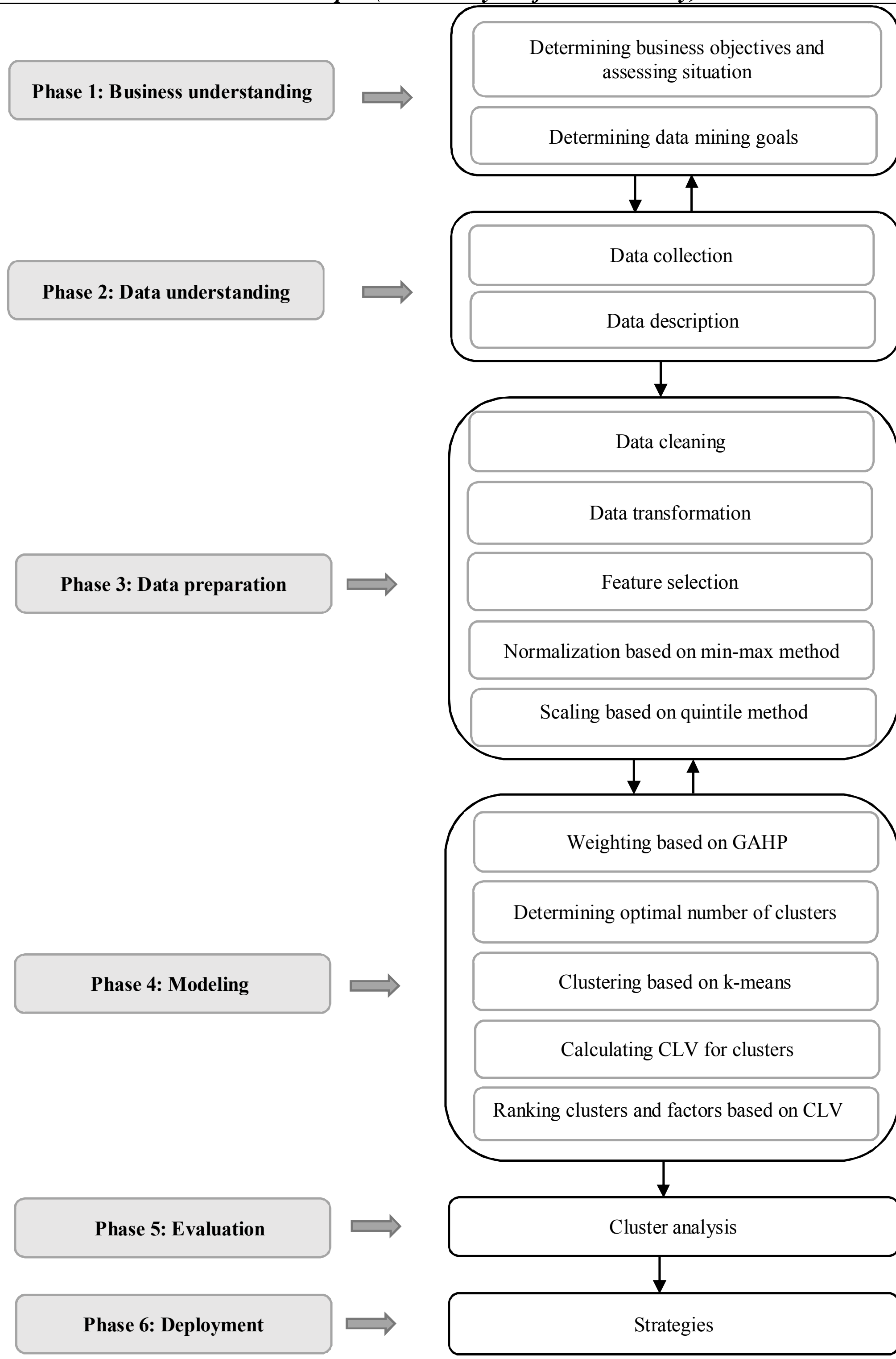

Fig. 3. Research roadmap

3.2.1. Business understanding

In the first phase, we are going to introduce the case study and describe its current situation. Besides, the objectives of running data mining on customers' data will be presented.

3.2.1.1. Determining business objectives and assessing situation

The company that is considered as the case study has been participating in software industry for 18 years producing thirty products used in various industries. Since this company has set an agenda for customer satisfaction, it has provided the customers with different products. Moreover, the needs of the customers have been correctly fulfilled by distinguishing their wants; this guarantees the value expected by the customers.

3.2.1.2. Determining data mining goals
Customer segmentation, identifying valuable customers, and providing the customers with value are the main goals of using data mining in this company.

3.2.2. Data understanding
In this step, data have been collected, identified, and finally described.

3.2.2.1. Data collection
The data set of customers includes a wide range of data which just demographic and transaction data are used and collected from 1865 customers from 2005 to 2018. The researchers of this study were provided with the data on January 20th, 2019.

3.2.2.2. Data description
The demographic data just include the provinces in which the company provides services; the customers belong to thirty one provinces all over Iran. Tehran has 24.8 percent of the customers (462 customers), 16.5 percent of customers (308 customers) belong to Esfahan and 8.2 percent of the customers (153 customers) belong to Kerman. As a result, Tehran, Esfahan and Kerman are the first, second and third provinces according to customer share.
As shown in Table 1, transaction data include the date of the first and last contract, the number of the contracts, and the amounts of money the customers have paid to the company based on Rial, the currency of Iran.

Tab. 1. Customers' transaction data

Customers	Date of the first contract	Date of the last contract	Number of contracts	Money paid to the company
Customer 1	2007/09/01	2018/11/03	43	31.551.254.296
Customer 2	2015/01/19	2018/10/13	12	26.344.000.000
Customer 3	2011/08/24	2018/12/11	9	16.641.300.000
Customer 4	2010/05/29	2018//09/12	18	10.999.460.000
Customer 5	2010/01/19	2018/08/15	12	9.585.948.000
.....	.....	.....	.....	.....
Customer 1865	2012/09/01	2014/08/04	2	1.000.000

3.2.3. Data preparation
Data preparation is considered as one of the most important and time-consuming phases in data mining. This phase includes 5 steps: data cleaning, data transformation, feature selection, normalization based on min-max method, and scaling based on quintile method. In this phase, the data collected from the previous phase is evaluated and then prepared for clustering.

3.2.3.1. Data cleaning
In the beginning, the data of 3 customers among 1868 customers were incomplete and therefore omitted from the customers' dataset. As a result, the data of 1865 active customers were evaluated.

3.2.3.2. Data transformation
Considering the transaction data of the previous phase and CLV model (including length, recency, frequency, and monetary), the primary data can be transformed to factors of CLV model (Table. 2). 20th January 2019 was the first day of study period in which the data was provided for the researcher. Finally, the data about the transformed factors which are new transaction data of customers have been presented in Table 3.

Tab. 2. Data transformation

Primary data	Factors
The number of days between the first and the last contract	Length
The number of days between the day of last contract and the first day of study period (2019/01/20)	Recency
Number of contracts	Frequency
Total money paid to the company by customers (Iranian Rial)	Monetary

Tab. 3. New transaction data of customers

Customers	Length	Recency	Frequency	Monetary
Customer 1	4081	78	43	31.551.254.296
Customer 2	1363	99	12	26.344.000.000
Customer 3	2666	40	9	16.641.300.000
Customer 4	3028	130	18	10.999.460.000
Customer 5	3130	158	12	9.585.948.000
.....	.....	.....	.....	.....
Customer 1865	702	1630	2	1.000.000

3.2.3.3. Feature selection

Selecting a collection of pertinent and important features and omitting unrelated features are two main purposes of feature selection algorithm [33]. Moreover, it is necessary to carry on this phase to select the correct factors before clustering [34]. To investigate the importance of CLV factors in this phase, feature selection algorithm has been used on SPSS Modeler 18. The inputs of the algorithm were length, rececny, frequency, and monetary. Running the algorithm, it was clearly shown that all those 4 factors were important and can be used in research model. Table 4 shows the results of the algorithm.

Tab. 4. Feature selection

Factors	Importance	Value
Length	Important	1
Recency	Important	1
Frequency	Important	1
Monetary	Important	1

3.2.3.4. Normalization based on min-max method

Normalization is defined as transforming all data into a common range from 0 to 1. One of the methods of normalization is min-max method. Based on Eq. (1), elements are normalized in a 0 to 1 continuum in which min_A and max_A are respectively the minimum and maximum of each element in each factor. Additionally, V and V* are primary amount and normalized amount of elements [35] , [36].

$$v* = \frac{v - min_A}{max_A - min_A} \quad (1)$$

As shown in Table 5, the maximum and minimum of each CLV factor have been extracted from customers' data. Table 6 indicates the amounts of data after normalization

Tab. 5. Minimum and maximum data

Factors	Min	Max
Length	0	4702
Recency	5	4494
Frequency	1	139
Monetary	1.000.000	31.551.254.296

Tab. 6. Normalized data

Customers	Normalized length	Normalized recency	Normalized frequency	Normalized monetary
Customer 1	0.867	0.983	0.304	1
Customer 2	0.289	0.979	0.079	0.834
Customer 3	0.566	0.992	0.057	0.527
Customer 4	0.643	0.972	0.123	0.348
Customer 5	0.665	0.965	0.079	0.303
.....	.....	.....	.....	.....
Customer 1865	0.149	0.638	0.007	0

3.2.3.5. Scaling based on quintile method

In LRFM model, the customers are evaluated based on the model's features. It's customary to use the real amounts of data to prove this common method. There is also another way that is widely used in some researches known as quintile method. In this method, the data are arranged descendingly.

At this point, the data are divided into 5 groups based on the number of customers or value of data. The top 20 percent of data are of the greatest value indicated by code (5). The next 20 percents of data are marked by code (4), (3), (2), and (1). In terms of value, code (5) is very high, code (4) is high, code (3) is medium, code (2) is low, and code (1) is very low [3] , [10]. Length, frequency, and monetary can be scaled correctly through quintile; however, due to the nature of rececny, counter scaling must be used. To run quintile, SPSS Statistics 25 has been used. Running binning and cut point 20%, the data were divided into 5 groups and scaled afterwards. Table 7 shows the steps of data division and scaling. Considering the new scales, Table 8 provides the data that will be further used to cluster the customers.

Tab. 7. Factor scaling

Scale	Length	Recency	Frequency	Monetary
Very high (5)	[1865-4702]	[5 -106]	[9-139]	[31.551.254.296 – 534.500.001]
High (4)	[1062-1864]	[107-210]	[5-8]	[534.500.000 – 246.200.001]
Medium (3)	[419-1061]	[211-397]	[3-4]	[246.200.000 – 126.940.001]
Low (2)	[1-418]	[398-1021]	2	[126.940.000– 70.450.001]
Very low (1)	0	[1022-4494]	1	[70.450.000 – 1.000.000]

Tab. 8. Research data based on quintile scaling

Customers	Length	Recency	Frequency	Monetary
Customer 1	5	5	5	5
Customer 2	4	5	5	5
Customer 3	5	5	5	5
Customer 4	5	4	5	5
Customer 5	5	4	5	5
.....	.....	.....	.....	.....
Customer 1865	3	1	2	1

3.2.4. Modeling

Modeling phase includes the following steps: weighting based on GAHP, determining optimal number of clusters, clustering based on k-means algorithm, calculating CLV for clusters, and ranking clusters and factors based on CLV.

3.2.4.1. Weighting based on GAHP

Analytic hierarchy process (AHP) is one of the best methods used in group multiple criteria decision making. In this method, a group of experts prioritized the factors based on a scale from 1 to 9, which was presented by Saaty in 1980 [37]. In fact, AHP is firmly based on pairwise comparison matrix (PCM) that's the basis for weighting factors in this research. To form this matrix, all factors are compared two by two [38] and eventually the final matrix is formed based on a geometric mean [39]. In this research, the relative importance of each factor has been determined by using group analytical hierarchy process (GAHP). The viewpoints of 17 experts were gained from pairwise comparison questionnaires and their reciprocal matrices were formed. Then, these reciprocal matrices are aggregated by using a geometric mean. Table 9 shows the final pairwise comparison matrix.

Tab. 9. Final pairwise comparison matrix

Factor	Length	Recency	Frequency	Monetary
Length	1	1.713	0.526	0.537
Recency	$\frac{1}{1.713}$	1	0.553	0.365
Frequency	$\frac{1}{0.526}$	$\frac{1}{0.553}$	1	0.325
Monetary	$\frac{1}{0.537}$	$\frac{1}{0.365}$	$\frac{1}{0.325}$	1

At this point, the elements of the final pairwise comparison matrix were put into expert choice software and the weighting process of factors was done that resulted in the information shown in Table 10. Clearly, monetary has the most weight (0.452) while rececny has the least weight (0.128). Inconsistency rate is 0.05, (less than 0.1), that shows the pairwise comparison matrix gained from the experts' views is adequately valid.

Tab. 10. Weights of factors

Factors	Length	Recency	Frequency	Monetary
Weight	0.186	0.128	0.234	0.452

3.2.4.2. Determining optimal number of clusters

Silhouette and SSE indexes have been used to determine the optimal number of clusters. Silhouette index considers the intercluster and intracluster distances at the same time and is between 1 and -1. The more this index is close to 1, the more clustering will be satisfying. To calculate this index, Eq. (2) is used in which S(x) is the amount of index, a(x) is average difference of x with the other data in the same cluster, b(x) is the minimum difference of x with the other data in other clusters [40].

$$\mathrm{s(x)} = \frac{\mathrm{b(x)-a(x)}}{\mathrm{max[b(x),a(x)]}} \qquad (2)$$

SSE index shows the sum of squared the distances of all data in each cluster with its center. This index is calculated through Eq. (3) in which X is data point in C_i cluster and m_i is a representative point for C_i cluster [41].

$$\mathrm{SSE} = \sum_{\mathrm{i=1}}^{\mathrm{k}} \sum_{\mathrm{x \in C_i}} \mathrm{dist}^2\,(m_i, x) \qquad (3)$$

To determine the optimal number of clusters, SSE and silhouette indexes for clusters have been calculated in Table 11. Silhouette index has been calculated through Matlab software. Moreover, SSE index has been calculated based on the SPSS Modeler outputs, which is gained from calculating
distance of each data from the cluster center and putting it in Eq. (3).

Tab. 11. Silhouette and SSE indexes for number of clusters from 2 to 10

Number of clusters	Silhouette	SSE
2	0.614	474.495
3	0.477	374.556
4	0.519	299.197
5	0.465	272.147
6	0.443	239.699
7	0.455	207.225
8	0.476	188.100
9	0.473	174.782
10	0.423	158.582

As shown in Fig. 4, a distinct knee in the SSE and a distinct peak in the silhouette index are present when the number of clusters is equal to 4. Silhouette index for number of cluster 4 is 0.519 and SSE index is 299.197. SSE index for the next number of clusters has a trend of decreasing, it has contained with a gentle slope after number of cluster 4 that can be disregarded. Thus, number of cluster 4 will be chosen as the optimal cluster.

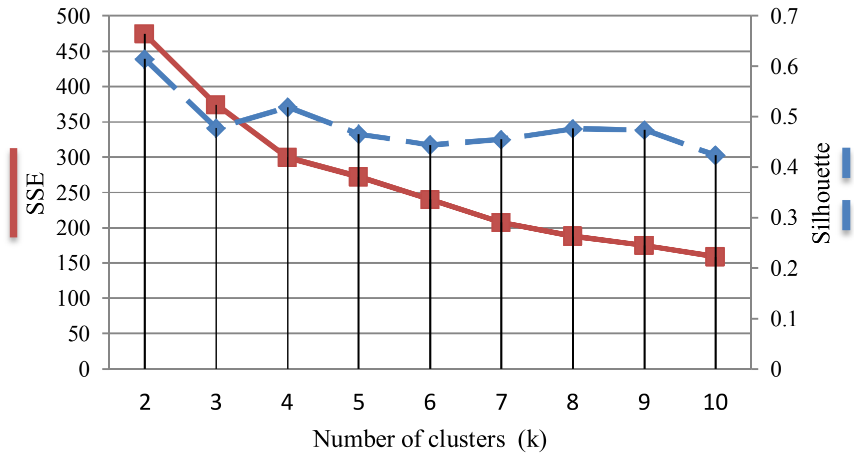

Fig. 4. SSE and Silhouette indexes for k from 2 to 10

3.2.4.3. Clustering based on k-means algorithm

Clustering the data is carried out on SPSS Modeler thorough k-means algorithm. As a result, the data have been divided in 4 clusters. While Cluster 1 has the most customers including 642 people and 34.42 percent of all the customers, Cluster 3 has the least customers including 301 people and 16.14 percent of all the customers (Table 12).

Tab. 12. The number and percentage of the customers in each cluster

	Cluster 1	Cluster 2	Cluster 3	Cluster 4
Number of customers	642	489	301	433
Percentage of customers	34.42	26.22	16.14	23.22

3.2.4.4. Calculating CLV for clusters

It is necessary to calculate the amount of CLV for each cluster through Eq. (4) in which: C^j indicates CLV for cluster j, W_L, W_R, W_F, and W_M show weight of each factor and C^j_L, C^j_R, C^j_F, and C^j_M are the average normalized LRFM amounts of cluster j [1]. The results are shown in Table 13.

$$C^j = W_L\, C^j_L + W_R\, C^j_R + W_F\, C^j_F + W_M\, C^j_M \quad (4)$$

Tab. 13. CLV calculation for each cluster

Factors	Weight	The average normalized LRFM			
		Cluster 1	Cluster 2	Cluster 3	Cluster 4
L	0.186	0.449	0.007	0.077	0.212
R	0.128	0.955	0.728	0.974	0.818
F	0.234	0.075	0.001	0.008	0.02
M	0.452	0.031	0.002	0.007	0.006
CLV	***	0.2373	0.0956	0.1440	0.1515

3.2.4.5. Ranking clusters and factors based on CLV

In this step, the clusters have been ranked using the results of the previous step. To do so, the cluster with the most CLV will be placed higher. Obviously, cluster 1 has the most CLV that is 0.2373 and the amounts of CLV for clusters 2, 3, and 4 are 0.1515, 0.1440, and 0.0956 respectively. Consequently,
cluster 1 will be ranked first and cluster 4, 3, and 2 will be ranked second, third, and fourth respectively.
To compare and rank the factors of CLV in each cluster, the average normalized LRFM in each cluster must be multiplied by their particular

weights. Table 14 indicates that the more a factor has value, the higher it will be ranked.

Tab. 14. Ranking factors in each cluster

Factors	Cluster 1		Cluster 2		Cluster 3		Cluster 4	
	Value	Rank	Value	Rank	Value	Rank	Value	Rank
L	0.083	2	0.001	2	0.014	2	0.039	2
R	0.122	1	0.093	1	0.124	1	0.104	1
F	0.017	3	0.0002	4	0.001	4	0.004	3
M	0.014	4	0.0009	3	0.003	3	0.002	4

3.2.5. Evaluation

In this phase, the customer clusters will be analyzed using the matrix presented in Fig. (1). The first step includes calculating the average amounts of CLV factors based on the primary data of all customers. Then, these amounts will be compared with the average amounts of CLV factors in each cluster, shown in Table 15.

Tab. 15. Analysis of the clusters

Factors	The average of total customers	The average of cluster 1	The average of cluster 2	The average of cluster 3	The average of cluster 4
L	1026.687	2113.557	34.413	362.308	1045.658
R	599.938	204.428	1221.987	119.86	817.581
F	5.468	11.362	1.198	2.189	3.831
M	449.675.504	991.033.820.70	95.635.629.45	238.035.933.20	193.964.118.90
LRFM	***	↑↓↑↑	↓↑↓↓	↓↓↓↓	↑↑↓↓

The results suggest that the customers of the first cluster are the high value loyal customers. Although these customers have greater length, frequency, and monetary than the average of the total customers, their rececny is lower than the average of the total customers.

The customers of the second cluster are uncertain lost customers with lower length, frequency, and monetary than the average of the total customers even thought their rececny is greater than the average of the total customers.

The third cluster includes the uncertain new customers. Considering these customers, the average of all the CLV factors is less than the average of the total customers.

Finally, the fourth cluster consists of the customers with high consumption cost who consume resources. While length and rececny of these customers are greater than the average of the total customers, frequency and monetary are less than the average of the total customers.

3.2.6. Deployment

In the last phase of Crisp-DM, the results gained from data mining and customer clustering have been applied to the processes of the software company so that the appropriate strategies for each cluster have been developed.

3.2.6.1. Strategies for cluster 1

This cluster has the most loyal and valuable customers with the highest CLV. Additionally, they bring huge amounts of profit for the company by repeatedly bulk purchasing. Considering the fact that these customers have established deep relationships with the software company for a long time, they are called 'loyal' customers. Therefore, customer maintenance measurements must be taken to boost customer loyalty. The best behavioral pattern for these customers is recognizing them as the most important customers of the company on whom honor must be bestowed. To do so, the company can hold seminars in which these customers are appreciated and relationships are improved. Moreover, loyal customer club may be a good idea to deliver them first-rate services. Customizing and presenting unique and high-quality products for these customers will definitely pay off.

Since loyal customers are essential for the company, they are entitled to receive not only particular discounts on the products, but also quick and high-quality services, proper after-sale services, on-time products notification, and the convenience of contacting the company more easily.

3.2.6.2. Strategies for cluster 2

Cluster 2 includes uncertain lost customers who haven't had long relationships with the company and haven't purchased any products recently. In fact, CLV of this cluster is not high. Besides, frequency and monetary of these customers are low. Consequently, they are not considered valuable for the company.
The first step toward these customers is finding out the reasons which make them break their relationships with the company. In this regard, a robust CRM system can pave the way for doing so by pointing out the reasons for customer attrition and the factors in customer dissatisfaction. Therefore, it's necessary to provide particular bonuses and attractive offers to lower customer churn.
Considering low CLV of the customers in this cluster and the fact that they don't create value, it's of great significance to state that this software company ought to investigate whether or not it's economical to bring back these customers. In fact before the customer churn occurs, the company should come up with a model based on the reasons of customer churn using data mining techniques and decision tree; this way, the company will be able to avoid customer churn and incorrect resource allocation to win back the lost customers.

3.2.6.3. Strategies for cluster 3

Cluster 3 has the least customers who are uncertain new. Since these customers have recently started purchasing and haven't had long relationships with the company, they are called new customers. The appropriate strategy toward these customers is to retain their cooperation with the company considering the fact that the cost of attracting a new customer is much more than retaining the current one. Furthermore, it's advisable to implement motivational plans such as holding repeatedly friendly meetings, on-time products notification, conducting constructive and consistent negotiations, and allocating special discounts. Another way to retain these customers is to profile and provide them with proper products in accordance with their needs through a CRM system. Lastly, a number of supportive methods such as information desk software and online dialog will culminate in success.

3.2.6.4. Strategies for cluster 4

Customers in cluster 4 have a high consumption cost. Since these customers have had long-lasting relationships with the company with little benefit over time, they are called resource consumer. Therefore, making new contracts with this cluster must be stopped, considering the fact that these customers are loss-making instead of value-making.
On the other hand, the company can start negotiations with these customers and offer them new products and services in order to encourage them to make new contracts by which not only the company gains considerable benefits, but the customers satisfy their needs. It's necessary to mention that these customers will change to core customers of the company provided that their frequency and monetary rise.

4. Conclusion

The concept of customer lifetime value has greatly enabled the software companies to cluster their customers, allocate their resources more appropriately, and design effective strategies to change low-value customers to high-value ones. Moreover, CLV plays a crucial role in preventing the organizations from wasting time, budget, and resources. In this regard, data mining has promoted CLV and made organizations develop and adopt marketing strategies, cut their expenditures, establish constructive relationships with their customers, improve the level of customer satisfaction and loyalty, and generate more income. Based on the research model and its gained results, the software company needs to design a behavioral pattern to treat its customers and implement appropriate policies to boost their values. In addition, it's necessary for the company to use decision tree and the techniques that results in extracting knowledge and discovering hidden patterns of each customer cluster.
The limitations of the study include difficulty in receiving customer data from the company, incoherent data of the departments, time-consuming nature of data integration, and lack of access to other demographic data.
For further research, it's highly recommended to use the research model in other branches of industry, use other clustering and data mining algorithms, and enhance LRFM model by adding more factors based on the business nature and the impact of these factors on CLV.

References

[1] Liu, D.-R. and Y.-Y. Shih, *Integrating AHP and data mining for product recommendation based on customer lifetime value.* Information & Management, Vol. 42, No. 3, (2005), pp. 387-400.

[2] Tsai, C.Y. and C.C. Chiu, *A purchase-based market segmentation methodology.* Expert Systems with Applications, Vol. 27, No. 2, (2004), pp. 265-276.

[3] Wei, J.-T., et al., *A case study of applying LRFM model in market segmentation of a children's dental clinic.* Expert Systems with Applications, Vol. 39, No. 5, (2012), pp. 5529-5533.

[4] Swarts, K.M., K. Lehman, and G.K. Lewis, *The use of social customer relationship management by building contractors: evidence from Tasmania.* Construction Management and Economics, Vol. 34, Nos. 4-5, (2016), pp. 302-316.

[5] Gupta, S. and D.R. Lehmann, *Customers as assets.* Journal of Interactive Marketing, Vol. 17, No. 1, (2003), pp. 9-24.

[6] Moro, S., P. Cortez, and P. Rita, *Using customer lifetime value and neural networks to improve the prediction of bank deposit subscription in telemarketing campaigns.* Neural Computing and Applications, Vol. 26, No. 1, (2015), pp. 131-139.

[7] Ferrentino, R., M.T. Cuomo, and C. Boniello, *On the customer lifetime value: a mathematical perspective.* Computational Management Science, Vol. 13, No. 4, (2016), pp. 521-539.

[8] Fu, X., et al., *User segmentation for retention management in online social games.* Decision Support Systems, Vol. 101, (2017), pp. 51-68.

[9] Akhondzadeh Noughabi, E. and A. Albadvi, *Mining the dominant patterns of customer shifts between segments by using top- k and distinguishing sequential rules.* Management Decision, Vol. 53, (2015), pp. 1976-2003.

[10] Peker, S., A. Koçyigit, and P.E. Eren, *LRFMP model for customer segmentation in the grocery retail industry: a case study.* Marketing Intelligence & Planning, Vol. 35, (2017), pp. 544-559.

[11] Swenson, E.R., N.D. Bastian, and H.B. Nembhard, *Data analytics in health promotion: Health market segmentation and classification of total joint replacement surgery patients.* Expert Systems with Applications, Vol. 60, (2016), pp. 118-129.

[12] Liu, D.-R., C.-H. Lai, and W.-J. Lee, *A hybrid of sequential rules and collaborative filtering for product recommendation.* Information Sciences, Vol. 179, No. 20, (2009), pp. 3505-3519.

[13] Däs, M., et al., *Customer lifetime network value: customer valuation in the context of network effects.* Electronic Markets, Vol. 27, No. 4, (2017), pp. 307-328.

[14] Estrella-Ramon, A., et al., *A model to improve management of banking customers.* Ind. Manag. Data Syst., Vol. 117, (2017), pp. 250-266.

[15] Gupta, S., et al., *Modeling Customer Lifetime Value.* Journal of Service Research - J SERV RES, (2006), Vol. 9, pp. 139-155.

[16] Hughes, A.M., *Boosting response with RFM.* Marketing Tools, Vol. 5, (1996), pp. 4-7.

[17] Chan, C.C.H., *Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer.* Expert Systems with Applications, Vol. 34, No. 4, (2008), pp. 2754-2762.

[18] Chang, H.H. and S.F. Tsay. *Integrating of SOM and K-mean in data mining clustering: An empirical study of CRM and profitability evaluation.* Journal of

Information Management, Vol. 11, (2004), pp. 161-203.

[19] Li, D.-C., W.-L. Dai, and W.-T. Tseng, *A two-stage clustering method to analyze customer characteristics to build discriminative customer management: A case of textile manufacturing business.* Expert Systems with Applications, Vol. 38, No. 6, (2011), pp. 7186-7191.

[20] Chen, S.C. and M.Y. Huang, *Constructing credit auditing and control & management model with data mining technique.* Expert Systems with Applications, Vol. 38, No. 5, (2011), pp. 5359-5365.

[21] Chen, C.-K., A.-J. Shie, and C.-H. Yu, *A customer-oriented organisational diagnostic model based on data mining of customer-complaint databases.* Expert Systems with Applications, Vol. 39, No. 1, (2012), pp. 786-792.

[22] Çakır, A., H. Çalış, and E.U. Küçüksille, *Data mining approach for supply unbalance detection in induction motor.* Expert Systems with Applications, Vol. 36, No. 9, (2009), pp. 11808-11813.

[23] Saltan, M., S. Terzi, and E.U. Küçüksille, *Backcalculation of pavement layer moduli and Poisson's ratio using data mining.* Expert Systems with Applications, Vol. 38, No. 3, (2011), pp. 2600-2608.

[24] Ait daoud, R., et al., *Clustering Prediction Techniques in Defining and Predicting Customers Defection: The Case of E-Commerce Context.* International Journal of Electrical and Computer Engineering, Vol. 8, (2018).

[25] Chiang, W.-Y., *Discovering customer value for marketing systems: an empirical case study.* International Journal of Production Research, Vol. 55, No. 17, (2017), pp. 5157-5167.

[26] Dursun, A. and M. Caber, *Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis.* Tourism Management Perspectives, Vol. 18, (2016), pp. 153-160.

[27] Liang, Y.-H., *Integration of data mining technologies to analyze customer value for the automotive maintenance industry.* Expert Systems with Applications, Vol. 37, No. 12, (2010), pp. 7489-7496.

[28] Cheng, C.-H. and Y.-S. Chen, *Classifying the segmentation of customer value via RFM model and RS theory.* Expert Systems with Applications, Vol. 36, (2009), pp. 4176-4184.

[29] Li, S.-T., L.-Y. Shue, and S.-F. Lee, *Enabling customer relationship management in ISP services through mining usage patterns.* Expert Systems with Applications, Vol. 30, No. 4, (2006), pp. 621-632.

[30] Chiang, L.-L. and C.-S. Yang, *Does country-of-origin brand personality generate retail customer lifetime value? A Big Data analytics approach.* Technological Forecasting and Social Change, Vol. 130, (2018), pp. 177-187.

[31] Sharma, S., K.-M. Osei-Bryson, and G.M. Kasper, *Evaluation of an integrated Knowledge Discovery and Data Mining process model.* Expert Syst. Appl., Vol. 39, (2012), pp. 11335-11348.

[32] Sen, B., E. Uçar, and D. Delen, *Predicting and analyzing secondary education placement-test scores: A data mining approach.* Expert Systems with Applications: An International Journal, Vol. 39, (2012), pp. 9468-9476.

[33] Li, H. and J. Sun, *Majority voting combination of multiple case-based reasoning of financial distress prediction.* Expert Systems with Applications, (2009), Vol. 36, pp. 4363-4373.

[34] Wang, C.-M. and Y.-F. Huang, *Evolutionary-based feature selection approaches with new criteria for data mining: A case study of credit approval data.* Expert Syst. Appl., Vol. 36, (2009), pp. 5900-5908.

[35] Venugopal, V. and S. Sundaram, *An online writer identification system using*

regression-based feature normalization and codebook descriptors. Expert Systems with Applications, Vol. 72, (2017), pp. 196-206.

[36] Jain, S., S. Shukla, and R. Wadhvani, *Dynamic selection of normalization techniques using data complexity measures.* Expert Systems with Applications, Vol. 106, (2018), pp. 252-262.

[37] Grošelj, P., et al., *Comparison of some aggregation techniques using group analytic hierarchy process.* Expert Systems with Applications, Vol. 42, No. 4, (2015), pp. 2198-2204.

[38] López-Ortega, O. and M.-A. Rosales, *An agent-oriented decision support system combining fuzzy clustering and the AHP.* Expert Systems with Applications, Vol. 38, No. 7, (2011), pp. 8275-8284.

[39] Collan, M., M. Fedrizzi, and P. Luukka, *A multi-expert system for ranking patents: An approach based on fuzzy pay-off distributions and a TOPSIS–AHP framework.* Expert Systems with Applications, (2013), Vol. 40, No. 12, pp. 4749-4759.

[40] Han, J., M. Kamber, and J. Pei, *10 - Cluster Analysis: Basic Concepts and Methods*, in *Data Mining (Third Edition)*, J. Han, M. Kamber, and J. Pei, Editors. Morgan Kaufmann: Boston. (2012), pp. 443-495.

[41] Ghousi, R., *Applying a decision support system for accident analysis by using data mining approach: A case study on one of the Iranian manufactures.* Journal of Industrial and Systems Engineering, Vol. 8, (2015), pp. 59-76.

Follow This Article at The Following Site:

Bakhshizadeh E, Aliasghari H, Noorossana R, Ghousi R. CUSTOMER CLUSTERING BASED ON FACTORS OF CUSTOMER LIFETIME VALUE WITH DATA MINING TECHNIQUE. IJIEPR. 2022; 33 (1) :1-16
URL: http://ijiepr.iust.ac.ir/article-1-1124-en.html