International Journal of New Political Economy 6(1): 23-46, 2025

Customer segmentation using the extended RFMP model based on Customer lifetime value (CLV) and data mining

Abdolah Saedi[1]*, Meysam Abbasi[2], Alieh Mehdi Nejad[3]

ARTICLE INFO

Article history:

Date of submission: 04 August 2024

Date of revise: 22 December 2024

Date of acceptance: 23 January 2025

JEL Classification:

M10, C15

C15, C15

C15

Keywords:

Customer relationship management (CRM), Clustering,

RFMP model,

Analytic hierarchy process (AHP),

K-means algorithm

ABSTRACT

One important way for identifying customers is clustering them to congruent segments. Skillfully clustering can be caused by identifying profitable customers by companies, understanding their requirement and allot their own resources in an appropriate way. We can also examine the change in the clustering of customers after the implementation of any economic policy and formulate and implement suitable public policies accordingly. The main goal of a recent study is clustering and identifying customers using of extended RFMP model. The study was done on 6665 customers of the Zahedan Refah chain store. According to this, after identification, the extended RFMP amounts (sum of recency, frequency, monetary and periodicity) weight of each variable was identified according to the analytic hierarchy process (AHP). At the next stage, optimal clusters number was specified by using Silhouette and Davies Bouldin indexes and customers were clustered with the K-means algorithm and finally, customers were preferred according to customer lifetime value. According to results, customers were divided into 3 parts and their traits were analyzed. Background can be provided for codifying relationship strategies with customers by using of recent study results.

1. Assistant Professor of Management Department. Faculty of Management and Economics, University of Lorestan, Khorramabad, Iran

2. Master of IT Management at the Department Of Management, Faculty of Management & Economic, University of Sistan and Baluchestan, Zahedan, Iran

3. Ph.D candidate at the Department Of Management, Faculty of Management and Economic, Sistan and Baluchestan University, Zahedan, Iran

* **Corresponding Author Email Address**: saedi.a@lu.ac.ir

DOI: https://doi.org/10.48308/jep.2025.236491.1183

Copyright: © 2025 by the authors. Submitted for possible open access publication under the terms and conditions of the Creative Commons Attribution (CC BY) license (https://creativecommons.org/licenses/by/4.0/).

1. Introduction

Nowadays, competing is an important issue for the business world and making a profit for a long time, attending to new customers is effective but the thing which should be concentrated is keeping old customers (Bagheri Ghalenooei and khorrami sarvestani, 2016). One of the effective strategies for retaining customers is customer relationship management (Rodrigez et al., 2018). CRM is a far-reaching corporative issue for understanding customer comportment, and influencing the long-time relationship with the customer, and improving customer loyalty, acquisition, retention and profitability (Shatnawi et al., 2017; Harwati & Karunia, 2017). CRM is used by many organizations, and those organizations that apply this approach should collect the huge amount of data and desiderata to get significant results from data (Shim et al., 2012). For this means, data mining can help CRM in different ways (Abu Ali and Abu Addose, 2010). Data mining is a method for trait's extraction patterns by using of large data set patterns which are not observed apparently and using different issues like clustering (Armano & Farmani, 2016; Azimi et al., 2017; Delgado et al., 2017; Sato et al., 2018), analyzing the principal ingredient (Li, 2014; Rezghi et al.,2014) and learning machine (Tahmasebi et al., 2017). Clustering is a kind of unconsidered learning mechanism, which is used for segmenting the customers by using behavioral data (Haevati & Adnan, 2017; Sato et al., 2018). Customer lifetime value (CLV) makes customer segmentation easy. The value which is existed in future profit circulation expected over an especial time for transacting with customers" is a definition for CLV (Safari et al., 2016). Customer lifetime value assessment is substantial because, for decision evaluating in the terrain of customer relationship management, it is used as a metric. For a firm, it is essential to make some anticipation for customers as each customer's lifetime value when the customer starts his/her occupations with the company, and at each purchase by the customer (Borle et al., 2008). One of the most important issues for evaluating customer lifetime value is RFM assessment (Liu & Shih, 2005). In RFM assessment customer, data is segmented by Recency (R), Frequency (F) and Monetary (M) variables (Dursun & Caber, 2016). By RFM admission, much profit can be obtained by the firm, embracing increase response rates, decreasing order cost and higher benefit can be reached (Liu and Shih, 2005). Because of the RFM model importance, some investigators try to progress towards this model and add some parameters to these three parameters (Khajvand et al., 2011). A recent study, the R*FMP model is used for customer segmentation of the Refah chain store in Zahedan. Refah chain store companies are activated as one of the most extent networks of distributing the modern-product in Iran with the goal of preparation and sales of basic products. For presenting suitable products to customers and supplying their requirements, customer segmentation is very substantial at this company. So, at this research, the customer was segmented by using the RFMP model, and these results display the context for codification customer relationship management strategies. In this model, the purchase period is examined. Because it is necessary to examine the purchase period as a variable. The innovation of the research is to examine the purchase period variable. This issue is also important for political economy. Since, it is necessary to examine the purchase period as a variable. The research innovation is to investigate of the purchase period variable. This issue is also important for political economy. Because in political economy, the study of economic systems (such as markets and national economies) and their management by political systems (such as law, institutions and the government) is addressed. Therefore, with the help of customer classification models, we can examine customer behaviors in relation to economic policies and formulate suitable administrative policies for managing them.

2. Theoretical framework and Literature Review

2.1. Customer relationship management (CRM)

Relationship Marketing (RM) principle is a suitable zone for modern-day marketing, and Customer Relationship Management is the root of that

(Rahimi and Kozak, 2017). Although the appearance of CRM as an important issue at a job and career is known there is no accepted definition for this (Ngai, 2009). Nowadays, CRM leads to a holistic issue for improving management of long-term profitable customer relationship and with this method, following the business processes and customer loyalty can be achieved (Pedron et al.2016). for creating a fantastic CRM in an organization all of the types of equipment and issues should be under controlled. CRM as a piece of equipment with high technology of Web/App makes an organization capable for comprehending customers or potential customers, and this matter helps the customers to make a suitable decision and creates good transactions. (Anshari et al., 2018). Some definition of CRM is offered. CRM is a commercial strategy and marketing that synthesis technology process and all activities of career (Anton & Hoeck, 2002). Engay emphasis that new definitions on the importance of CRM as a strategic process for the maximum value of a customer for the organization. Ballantyne 2003 the concept of CRM has related to the development of an organization that requirements and preferences of the customer had another definition (ballantyne, 2003). Emerson, 2004 defines that the basis of CRM is relationship marketing logic and codification of goals and strategies of that. CRM is meaning a work to manage the relationship management on a big scale with long profit persistence in the customer's mind. CRM is defined as an intensive career strategy on customer and its goal is increasing the customer's satisfaction and loyalty (Seeman & hara, 2006). Nowadays, marketing managers distinguish is concentrated on CRM on stale and long relationships with the customer, and it is valuable for both sides (Nguyen, 2007).

2.2. Data mining

The process which is using machine learning and other methods for classical databases in order to extract implicit, previously unknown and potentially useful patterns from a database is a definition for data mining. In knowledge exploration, data mining can be considered as a substantial step (Khajevand

& Jafar Tarokh, 2011). This process is performed by combining statistical analysis and machine learning for the hidden pattern's extraction and relationship from large data sets. As it is mentioned before, data mining is known as an essential step in knowledge explorations. It is large datasets explorations for discovering hidden and previously unknown patterns, relationships and the knowledge which is complicated to discover with traditional statistical approaches (Shamsollahi et al., 2019). Data mining can be segmented into the following stages:

- relationship's determination and regularities are defined as the first stage;
- applying the rules to forecast other objects is known as the second stage;
- Variance analysis is expressed as the third stage. This process assesses deviation and error analysis. So, data mining can be represented as a sequence of three steps.

Explaining the relationships – using these legislations – analyzing the variance (Vadim, 2018).

The volume of data contained in databases has grown quickly as a result of ongoing development of database technology (Xianget al., 2024), and the extensive adoption of database management systems. Behind these huge data, there is a large amount of valuable information hidden. If this information can be extracted from the database, enormous benefits can be brought. This technology that extracts information from large-scale databases is called data mining. Data mining, in its broadest sense, is the process of obtaining hidden and potentially valuable knowledge and information from vast, imperfect, noisy, fuzzily distributed, and random data. The definition contains the following four meanings: (1) the data source must be real, massive and noisy. (2) What is found is meaningful knowledge for the user. (3) The knowledge obtained must be acceptable, understandable, and applicable, and the results should preferably be expressed in a natural language (lue et al., 2023). This does not mean to find a general knowledge or a new, pure and scientific theory, but only a relatively necessary condition and limitation to apply to certain fields of theoretical knowledge. From a business perspective, data mining can be defined as an advanced and effective way to discover and model hidden, unknown or known laws by exploring and analyzing a large amount of enterprise data based on business goals set by an enterprise.

2.3. Customer segmentation

Customer segmentation is seen as one of the pillars of a successful advertising campaign (Moulay & Hanaa, 2023).

Customer relationship management (CRM) at supermarkets is willing to interact with customers appropriately with the aim of making strong relationship and resultantly gaining maximum profits. Customers consist of various groups of people and have different needs, styles and expectations (Maraghi et al., 2020).One of the decisive marketing strategies with the goal of market segment's identification and designation, which can be a main purpose for the firms' marketing plan, is known as customer segmentation. Segmentation is essential if the main target of industry employers is to obtain customer satisfaction that can maintain the customers (Gichuru, M. J. & Limiri, E. K, 2017). By identifying the differences among customers, they can be segmented into different groups, and it is the main purpose of customer segmentation. By segmenting the customers some information like: customer demographics (age, race, religion, gender, family members, ethnicity, income and education level), geographic (where they live and work), psychographic (social class, lifestyle and personality traits) and behavioral (spending, consumption, usage and desired benefits) can be reached. Customer segmentation can influence on customer service and develop that and customer maintenance, and loyalty can be created and improved. As a by-product of its personalized nature, marketing materials sent out using customer segmentation tend to be more valued and appreciated by the customer who receives them as impersonal brand messaging that doesn't acknowledge purchase history or any kind of customer relationship. Customer management can be influenced by customer segmentation; by segmenting the customers into different groups, that share similar requirements, the firm can concentrate on each segment and their requirements. According to the firm's resources or requirement's large or small segments can be disported (Melnic, 2016).

2.4. Extended RFMP model

RFM is a kind of model, which is differed from remarkable customers from the data lump by three epithets. These three facts ures are: customer consumption recency, frequency and value of Monetary. The RFM details are expressed as follows (Cheng & Chen, 2009):

1. The recency of the last purchase which is observed by (R) refers to the period of time between last purchase time and certain scrutiny period. A lower interval shows the higher value of this variable in the model.
2. F represented purchase frequency and it means the transactions number at a certain period of time. When the purchase repetition is higher, the optimal situation shows bigger (F).
3. M, which is represented monetary value refers to the value of money consumption by the customer in a special period of time.

According to the basic model of RFM, at recent research R*FMP model is used. And it is expressed as a formula as follows:

R*: recency factor is defined as collected recency of all transactions which is done by the customer, this recency factor has three advantages than recency definition at the reference model; first: the last transaction of customer isn't attended, but all customer transactions are considered, second: when the last moment of the special period is passing, the attendances of that time are more considerable than the beginning, if the study period is passed, recency factor goes higher, and it is calculated as below formula:

$$R^* = \sum_{i=1}^{f}(D_i - D_0) \tag{1}$$

In the mentioned formula, f shows the number of transactions, which are done by the customer, do represent the date which is referred to the beginning of the study. Di represented the i purchase of customer, and i is counter for customer transactions. (Gholamian & Niknam, 2012).Periodicity: This new feature reflects whether customers visit the stores regularly. We define periodicity as the standard deviation of the customer's inter-visit times:

$$Periodicity = stdev(IVT_1, IVT_2, \dots, IVT_{n-1}, IVT_n) \quad (2)$$

Where IVT denotes the inter-visit time and n represents a customers number of inter-visit time values. IVT is the elapsed time in days between two consecutive visits of the customer, and it is defined as follows:

Where i≥1 and ti denote the date corresponding to ithvisit of the customer (Peker et al, 2017).

Customer existence at regular intervals is an appropriate definition for periodicity. Low periodicity value showed that the aforementioned customer exists or makes purchases at comparatively fixed intervals and can be known as regular.

2.5. K-means

In the field of segmentation, K-means is iterative. It is a standard algorithm in that it takes, as input, the data and the number of clusters ‘‘k”. The output is data divided into ‘‘k” clusters so that the resulting intra-cluster similarity is high (minimizing the sum of squares within clusters in equation (Smaili & Hachimi, 2020). K-means algorithm that based on Euclidean distance function clustering is the process of segmenting a set of physical or abstract objects into the segments of similar traits. The set of data with the same characteristics form a cluster which is different from other clusters (Han & Kamber, 2001). One of the famous and popular clustering algorithms is K-means, and it is known as Forgy's method (Forgy, 1965), and it is applied at different parts like data mining, statistical data analysis, and other business application. So, the main goal of the recent study is making the clusters according to the same traits. K-means algorithm segmentation is based on the mean value of res in the cluster. K-means is a very susceptible selection of beginning point for dividing the members into K initial clusters. Performance comparison of variable clustering strategies applied intraclass method when the fixed cluster number of K value is explained as follows (Seyed hosseini, 2010):

$$F(k) = \frac{1}{k \sum_{n=1}^{k} \sum Dist(c_i, c^n)} \quad (3)$$

According to the formula the computing process for K-means is defined as:

Step 1: division the members into K initial clusters. Dividing the members (m objects) into K initial clusters. Step 2: put a member into a segment that has the nearest traits (the distance is calculated by using of Euclidean distance with prior observations), and the priority is computed again when some observations are lost or added. Step 3: repeating step 2 would be done until no more reassigning. Rather than starting with division all members into K introductory segments at step 1, specifying K initial seed points and then proceed to step two should be done. Some items which are dependent on initial division or initial selection of seed points can be known as the last mission. Experiences offer that the most main changes created when the first redistribution step is done (Mirzaean Rajeh et al., 2014).

2.6. Literature Review

Moulay & Hanaa (2023) have modified the RFM model by adding diversity ‘‘D” as a fourth parameter, referring to the diversification of products purchased by a given customer. The segmentation based on RFM-D is applied in a retail market in order to detect behavior patterns for a customer. The proposed model increases the quality of prediction of customer behavior; Companies could predict, customers who will respond positively. Christy et al (2021) initially perform an RFM analysis on the transactional data and then extend to cluster the same using traditional K-means and Fuzzy C- Means algorithms. In this paper, a novel idea for choosing the initial centroids in K- Means is proposed. The results obtained from the methodologies are compared with one another by their iterations, cluster compactness and execution time. Maraghi et al (2020) present a model completing CRM process from understanding customers to assigning marketing strategies. Profitable customers will be distanced as a result of correct understanding of all customers. Research is comprised of two phases. At phase one, dataset with recency, frequency and monetary (RFM) measures is constructed and clustered using K-means algorithm. Six segments of customers are detected based on the results of clustering. All segments are comprehensively analyzed and marketing strategies for them are described in phase two. Transactions of every segment of customers are separated and association rules are extracted using market basket analysis and Apriori algorithm. Consequent and also antecedent product items are proposed to customers who purchased antecedent product iems. So, dedicated marketing proposals are developed for some special customers. Imani and Abbasi (2017) analyzed customers of a chain store in Iran using the RFM model. They used the Fuzzy analytic hierarchy process and fuzzy C- FCM algorithm, and customers were divided into seven clusters that provided a framework for developing customer relationship strategies (Imani and Abbasi, 2017). Peker et al (2017) in research analyzed customers of a grocery company. Using the lrfmp model and the k-means algorithm and divided customers into 5 clusters (Peker et al., 2017). Safari et al (2016) in research analyzed customers of a store. They using the fcm algorithm divided customers into 9 clusters and ranked clusters by calculating CLV (Safari et al., 2017). Tleis et al 2017 cluster organic food market in their study at Lebanon. For doing this research 320 questionnaires between consumers of organic food at Beirut was distributed. By analyzing the questionnaire, consumers by using K- means algorithm was clustered to 4 clusters and for each cluster, suitable strategies were codified (Tleis et al., 2017). Hiziroglu and Senbas 2016 at their research, with using of the Fuzzy clustering automotive industry of customer. The aim of this study was comparing that with Fuzzy clustering by using the algorithm of the traditional segmentation method. Clustering, for this research the data of 130 customers of the automotive supplier was received. Results showed more balanced clustering than the traditional one and help the marketing managers to understand their customers more (Hiziroghlu & Senbas, 2016).

4. Methodology

At recent research one of the largest chain stores in Iran was clustered according to variables of the R*FMP model and by using of K-means algorithm. At this study, stages of data mining and data analysis were performed for discovering the knowledge of them according to the standard process of CRISP-DM. this process includes System understanding, Data's understanding, Data preparation, Modeling, Model evaluation, and Deployment.

5. Analysis of results

Step one: System understanding

The Refah chain store as one of the most extents of modern product distribution networks in the country has the goal of preparation, supplying, distribution and sales of basic products. This firm tries to give its products to customers by holding sales and seasonal festivals; and according to sales data registration increment at firms and applying them by data mining knowledge experts, customer value identification can be effective for the managers of this kind of firm. The main goal of the recent investigation is considering the sum of purchase recency variable and Periodicity purchase order variable to the RFM model, so by this explanation's customer analysis can be done.

Step two: Data understanding

The preparation data stage contains all activities which are applied for composing ultimate data collection (data that are prepared for modeling) from raw primary data. If the preparation quality was higher, modeling will be better too. According to these subjects, at this stage, incomplete and wrong data, invalid amounts and some transactions with lost amounts were eliminated.

Step Three: Data preparation

The preparation data stage contains all activities which are applied for composing ultimate data collection (data that are prepared for modeling) from raw primary data. If the preparation quality was higher, modeling will be better too. According to these subjects, at this stage, incomplete and wrong data, invalid amounts and some transactions with lost amounts were eliminated. Applied variables at this study include:

Table 1. Calculate the values of R*FMP model variables

customer number	R*	F	M	P
1	83	4	1258609	9.8149
2	74	3	1281680	11.3137
3	143	6	7232771	3.2093
4	164	7	3762029	4.2426
5	451	11	1844536	6.2503
. . .	. . .	. . .	. . .	. . .
6665	153	3	1537800	42.4264

Source: research findings

Because of differences at the RFM model variable's unit, the amount of these variables should be normalized according to an identical unit. For this purpose, the minimum-max normalization method is used until all numbers collected between 0 to1. The bellow equation shows normalization by the minimum-max method.

$$NR = \frac{R_{max}-R}{R_{max}-R_{min}} \quad NF = \frac{F-F_{min}}{F_{max}-F_{min}} \quad NM = \frac{M-R_{min}}{M_{max}-M_{min}}$$

$$NP = \frac{P_{max}-P}{P_{max}-P_{min}} \tag{4}$$

In the above relationships, Rmax, Fmax, Mmax and Pmax represent the highest values of the variables, and Rmin, Fmin, Mmin and Pmin also represent the lowest variable values. R, F, M and P also show the main values of the variables. Finally, NR, NF, NM and NP also represent normalized values of the variable.

Table 2. Normalized values of each of the R*FMP model variables

Customer number	NR*	NF	NM	NP
1	0.008	0.0137	0.0102	0.9121
2	0.0071	0	0.0104	0.8987
3	0.0146	0.0411	0.0643	0.9713
4	0.0169	0.0548	0.0328	0.962
5	0.0481	0.1096	0.0155	0.944
. . .	. . .	. . .	. . .	. . .
6665	0.0157	0	0.0127	0.6202

Source: research findings

With using of RFMP, weight relevant to the model variable is calculated by the analytic hierarchy process approach and includes three steps. At first, step, (AHP), experts and decision-makers should perform pairwise among variables. At the second step, the compatibility of pairwise should be assessed according to the incompatibility index. The incompatibility index is accepted when the amount of that is lower than 0.1. At the third step, the variable's weights are calculated. For performing this stage, I wanted the interior manager and five experts of the sales units to perform pairwise between variables: recency, frequency, Purchase periodicity and monetary. In continues, for analyzing the comparison compatibility, the variable's weight calculation software was used. Comparison compatibility was 0.02, and it was accepted. Obtained weights for RFMP variables are collected in table 3.

Table 3. R*FMP model variables weight

Variables name RFMP	Obtained weights
Recency	0.178
Frequency	0.439
Monetary	0.265
Purchase periodicity	0.11

Source: research findings

The table below shows the normal and weighted values of the model variables.

Table 4. Normal and weighted value of R*FMP model variables

Customer number	NR*×WR*	NF×WF	NM×WM	NP×WP
١	0.0015	0.006	0.0027	0.1003
2	0.0013	0	0.0027	0.0989
3	0.0027	0.018	0.017	0.1068
4	0.0032	0.024	0.0087	0.1059
5	0.009	0.0481	0.0041	0.1038
. . .	. . .	. . .	. . .	. . .
6665	0.0029	0	0.0033	0.0682

Source: research findings

Step Four. Modeling

A recent study K-means algorithm was used for customers clustering. For this means SPSS Modeler 18 software was used, and it is one of the most famous software for data mining. Clustering around quality assessment, the superiority of one cluster toward other clusters is measured by different clustering algorithm's same algorithms but with different parameter's amount, and at this study, two indexes of Silhouette and Davies Bouldin were used. Silhouette segregation and density are shown with weak, medium and good. The mean of the Silhouette index amount is applied for clustering validity assessment decision making for choosing optimal classes, and this amount is calculated according to remoteness and proximity of observations and clusters to each other. The amount of this index is different from -1 to +1. If it is near to +1, it means that sample clustering was well done, and suggested cluster is appropriate, and if it is near to -1, it means that sample clustering wasn’t suitable and suggested cluster for special data isn't proper. Davies-Bouldin Index calculates the similarity mean of each cluster observation, and for assessment, the lake of similarity between clusters is used. If the amount of this index is lower, the clustering quality is higher. According to the expert's suggestions and stores, managers cluster number should be between 3 to 10. The below table shows Silhouette and Davies Bouldin indexes amount for 3 To 10 clusters.

Table 5. Silhouette and Davies Bouldin indices amount

number of cluster	Davies Bouldin	Silhouette
3	0.758	0.68
4	0.83	0.532
5	0.765	0.532
6	0.76	0.53
7	0.888	0.517
8	0.881	0.521
9	0.903	0.521
10	0.942	0.51

Source: research findings

According to the table results, when customers divide to three clusters, the amount of Silhouette and Davies Bouldin indexes are maximum and minimum respectively, so this number (three clusters) is chosen as the optimal number of clusters for this study. In continuation, details of clustering results with K-means algorithms in the format of table number 6 are shown.

Table 6. Details of clustering results with K-means algorithm

Cluster	Average of R*	Average of F	Average of M	Average of P	Number	Percentage
1	0.0066	0.0036	0.0084	0.0604	1469	22%
2	0.0722	0.1972	0.0501	0.1044	108	1.6%
3	0.0106	0.0156	0.0114	0.0955	5088	76.3%
Total average	0.0108	0.0159	0.0114	0.0879		

Source: research findings

Step Five: Model evaluation

The clustered analysis is performed from a comparison of R*FMP model variables mean at each cluster and that variable means at all data and also ranking the variables according to the average amount of R*FMP model variables.

Table 7. Clustered analysis of data

Cluster	Average situation of variables (R, F , M , p)	Average rank of variables (R, F , M , p)
1	(↓ , ↓ , ↓, ↓)	(3 , 3 , 3 , 3)
2	(↑ , ↑ , ↑, ↑)	(1 , 1 , 1 , 1)
3	(↓ , ↓ , =, ↑)	(2 , 2 , 2 , 2)

Source: research findings

At table six mean of each suggested R*FMP model variable at each cluster is compared with the mean of that for all data. At this table, sign ↑shows a situation that means of the variable at one cluster is more than the value means of that variable at all data, and it represents the optimal situation of that variable. The sign ↓ shows that variable mean at one cluster is lower than the mean of the mentioned variable at all data, and it defines an undesirable situation of that variable. As it is indicated, at second cluster amounts of all research model variables are higher than the total mean of data, and it shows a desirable situation. At the third cluster, the sum of recency (R*) and frequency (F) is lower than the total mean of data, but the periodicity (P) variable is higher than the total mean. At first cluster, three variables of R*, M & P are lower than the data mean, but the F variable is higher than the total mean of data. At this stage Customer Lifetime, Value (CLV) is identified. CLV is a core metric in customer relationship management. It can be useful to improve market segmentation and resource allocation, evaluate competitor firms, customize marketing communication, optimize the timing of product offerings, and determine a firm's market value (Wirawan Dony Dahana, Yukihiro Miwa, Makoto Morisada, 2019). From the sum of 4 variables value, recency, frequency, purchase monetary and purchase periodicity, CLV of each cluster is obtained and ranking them will be performed.

Table 8. Calculation of CLV and cluster ranking

cluster	CLV	Rank
1	0.079	3
2	0.4239	1
3	0.1331	2

Source: research findings

At this pyramid, CLV at three clusters is defined; at the left part of this pyramid, the identified numerical amount for CLV of each cluster and at the right part of that customer number of that cluster is written. Anyway, moving from low level to high level is performed, confronting to customers with higher CLV is shown.

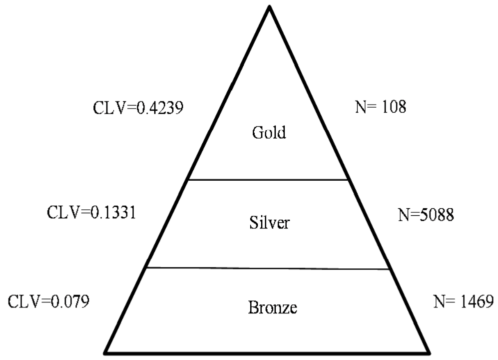

Figure 1. Customer value pyramid (Source: research findings)

Step Six: Deployment

In this stage, the final report from ex-stages results will express. Customers of segment 1: This segment contains the least value of store customers; they form 22% of all customers, and their CLV is 0.079. So, they take place at the third rank and at the end of the pyramid. This segment is named a bronze segment. The whole R*FMP variables average is lower than the total mean of data. So, for making relationship programs with the customer, an appropriate balance between costs or relationship with the customer with their profitability for store must be considered. Therefore, those kinds of customers should be identified and maintained that can be valuable customers because keeping all the customers has a high cost for the store. Customers of segment 2: This segment contains the most valuable customers. All of the R*FMP model variables are higher than total mean, and they are known as golden customers and they take place at the head of the pyramid. It almost 2% of all store customers and their CLV is 0.4239. According to this, the store must try to keep them; and by codifying appropriate loyalty strategies must try to increase their satisfaction. Customers of segment 3: This segment form 76.3% of all store customers. Their CLV is 0.1331, and they are at the second rank and middle of the pyramid, and they are known as the silver group. R*, F & M variables are the same as total mean and only the P variable is higher than the total mean. According to the purchase regularity of this segment, sales programs and seasonal sales festivals must be codified to improve R*FM variables; because the customer's number of this segment is high and good capacity for transforming to golden segment is existed.

6. Conclusions and policy suggestions

One of the most important factors for persistence at dynamic markets is understanding and adapting with occurred changes at customer's behavior. In this article, customers of one chain store are clustered by using data mining techniques. According to this context that, at the basic model of RFM, recency depends on the last transaction. Recent investigation tries to use the variable of the sum of customer’s recency. And also, the P variable is used as a priority for customers to purchase discipline measurement. After identifying the amount of the R*FMP variable, the analytic hierarchy process (AHP) was used for their weight's identification. According to clustering with K-means algorithm results, customers divided into three clusters. After calculating CLV for each cluster, their ranks were identified, and they have known as golden, silver and bronze.We can also examine the change in the clustering of customers after the implementation of any economic policy and formulate and implement suitable public policies accordingly. These results help the store to represented appropriate services for customers according to their segment and eliminated to apply traditional approaches for all customers. It is suggested to use this variable at future investigations, and its results must be compared with the primary model of RFM. And it is also suggested to use the other clustering algorithms and by completing customer population cognition information like age, sex, education and marital status more complete programs of relationship with the customer can be represented. The following policy suggestions are also provided:

- Using strategies that turn silver and bronze customers into gold.
- Knowing the needs of bronze and silver customers

- Knowing the priorities of bronze and silver customers
- Understanding the behavior of competitors Appropriate and timely response to market changes
- Predicting customer behavior

Funding

This study received no financial support from any organization.

Authors' contributions

All authors had contribution in preparing this paper.

Conflicts of interest

The authors declare no conflict of interest

Reference

AbuAli, A. N. & Abu Addose, H. Y. (2010). Data Warehouse Critical Success Factors. *Scientific Research*, 42(2), 326-335.

Anshari, M., Almunawar, Ariff Lim, M. N., & Syamimi, M, A. (2018). Customer relationship management and big data enabled: Personalization & customization of services. *Applied Computing and Informatics*, 1-8.

Anton, j. & Hoeck, m. (2002). *E-business customer service*. Santa Monica, the Anton press.

Armano, G., & Farmani, M. R. (2016). Multiobjective clustering analysis using particle swarm optimization. *Expert Systems with Applications*, 55, 184–193.

Azimi, R., Ghayekhloo, M., Ghofrani, M., & Sajedi, H. (2017). A novel clustering algorithm based on data transformation approaches. *Expert Systems with Applications*, 76, 59–70. (in persian)

Bagheri Ghalenooie, M. & Khorrami Sarvestani, H. (2016). Evaluating

Human Factors in Customer Relationship Management Case Study: Private Banks of Shiraz City. *Procedia Economics and Finance*, 36, 369-373. (in persian)

Ballantyne, D. (2003). A relationship-mediated theory of internal marketing. *European Journal of marketing*, 37(9), 1242-1260.

Borle, S., Singh, S. S., & Jain, D. C. (2008). Customer Lifetime Value Measurement. *Management Science*, 54(1), 100–112.

Cheng, C. H., & Chen, Y. S. (2009). Classifying the segmentation of customer value via RFM model and RS theory. *Expert systems with applications*, 36(3), 4176-4184.

Christy, A., J., Umamakeswari, A., Priyatharsini, L., & Neyaa, A. (2021). RFM ranking – An effective approach to customer segmentation. *Journal of King Saud University Computer and Information Sciences,* 33, 1251- 1257.

Delgado, S., Higuera, C., Calle-Espinosa, J., Mor´an, F., & Montero, F. (2017). A SOM prototype-based cluster analysis methodology. *Expert Systems with Applications*, 88, 14–28.

Dursun, A., & Caber, M. (2016). Using data mining techniques for profiling profitable hotel customers: An application of RFM analysis. *Tourism Management Perspectives*, 18, 153–160.

Gholamian, m., Niknam, z. (2012). Provide an adapted model for segmentation of banks' customers based on their lifetime value. *Journal of Executive Management*, 4(7), 5-5. (in persian)

Gichuru, M. J. & Limiri, E. K. (2017). Market Segmentation as a Strategy for Customer Satisfaction and Retention. Economics. *Commerce and Management*, 7(12), 544-553.

Han, J., & Kamber, M., (2001). Data mining: Concepts and techniques. San Francisco: *Morgan Kaufmann Publishers*.

Harwati, A., M. & Karunia, A. (2017). Utilization of Social for Consumer Behavior Clustering using Text Mining Method. *Engineering and Applied Sciences*, 12(3), 6406-6411.

Hiziroglu, A., & Senbas, U. D. (2016). An Application of Fuzzy Clustering to Customer Portfolio Analysis in Automotive Industry. *Fuzzy System Applications (IJFSA)*, 5(2), 13-25.

Khajvand, M & Jafar Tarokh, M. (2011). Analyzing Customer Segmentation Based on Customer Value Components (Case Study: A Private Bank: Technical note). *Industrial Engineering, Special Issue*, 79-93. (in persian).

Li, H. (2014). Asynchronism-based principal component analysis for time series data mining. *Expert Systems with Applications*, 41, 2842–2850.

Liu, D.-R., & Shih, Y. Y. (2005). Integrating AHP and data mining for product recommendation based on customer lifetime value. *Information & Management*, 42(3), 387–400.

Luo, Y., Yu, X., Yang, D. (2023). A survey of intelligent transmission line inspection based on unmanned aerial vehicle, Artif. Intell. Rev. 56 (1), 173–201.

Maraghi, M., Adibi, M. A., & Mehdizadeh, E. (2020). Using RFM Model and Market Basket Analysis for Segmenting Customers and Assigning Marketing Strategies to Resulted Segments. *Journal of Applied Intelligent Systems and Information Sciences*, *1*(1), 35-43. (in persian)

Melnic, E. L. (2016). How to strengthen customer loyalty, using customer segmentation? *Economic Sciences*, 9(58), 2, 52-60.

Mirzaiean Rajeh, Sh., Adeli Koudehi, F.; Seyedhosseini, S. M. & Farazmand, R. (2014). A Model for Customer Segmentation Based On Loyalty Using Data Mining Approach and Fuzzy Concept in Iranian Bank. *Business and Behavioral Sciences*, 4(9), 118-136. (in persian)

Moulay, Y., S., & Hanaa, H. (2023). New RFM-D classification model for improving customer analysis and response prediction. *Ain Shams Engineering Journal,* 14(12).

Ngai, E. W. T., Xiu, L., & Chau, D. C. K. (2009). Application of data mining techniques in customer relationship management: A literature review and classification. *Expert Systems with Applications*, 36, 2592–2602.

Nguyen, T. H., Sherif, J. S., & Newby, M. (2007). Strategies for successful CRM implementation. *Information Management & Computer Security*, 15(2), 102-115.

Peker, S., Kocyigit, A., & Eren, P. E. (2017). LRFMP model for customer segmentation in the grocery retail industry: A case study. *Marketing Intelligence & Planning*, 35(4), 544-559.

Rahimi, R., & Kozak, M. (2017). Impact of customer relationship management on customer satisfaction: The case of a Budget Hotel Chain. *Travel & Tourism Marketing*, 34(1), 40–51. (in persian)

Rodriguez, M., Peterson, R. M., & Krishnan, V. (2018). Impact of CRM technology on sales process behaviors: empirical results from US, Europe, and Asia. *Business-to-Business Marketing*, 25(1), 1–10.

Safari, F., Safari, N., & Montazer, G. A. (2016). Customer lifetime value determination based on RFM model. *Marketing Intelligence & Planning*, 34(4), 446–461. (in persian)

Sato, Y., Izui, K., Yamada, T., & Nishiwaki, S. (2018). Data mining based on clustering and association rule analysis for knowledge discovery in multi objective topology optimization. *Expert Systems with Applications*.

Smaili, M. Y., & Hachimi, H. (2020). Hybridization of Improved Binary Bat Algorithm for Optimizing Targeted Offers Problem in Direct Marketing Campaigns. Advances in Science, Technology and Engineering Systems Journal. 5(6):239–46. (in persian)

Seeman, E. D., & O'Hara, M. (2006). Customer relationship management in higher education: Using information systems to improve the student-school relationship. *Campus-Wide Information Systems*, 23(1), 24-34.

Seyed Hosseini, S. M., Maleki, A., & Gholamian, M. R. (2010).Cluster analysis using data mining approach to develop CRM methodology to assess the customer loyalty. *Expert Systems with Applications*, 37, 5259–5264. (in persian)

Shamsollahi, M., Badiee, A. & Ghazanfari, M. (2018). Using Combined Descriptive and Predictive Methods of Data Mining for Coronary Artery

Disease Prediction: a Case Study Approach. *AIand Data Mining*, 7(1), 47-58. (in persian)

Shatnawi, M. Q., Yassein, M. B., & Al-natour, H. (2017). Customer relationship management at Jordan University of science and technology: Case study, issues and recommendations. *International Conference on Engineering and Technology (ICET)*.

Shim, B., Choi, K., & Suh, Y. (2012). CRM strategies for a small-sized online shopping mall based on association rules and sequential patterns. *Expert Systems with Applications*, 39(9), 7736–7742.

Tahmasebi, P., Javadpour, F., & Sahimi, M. (2017). Data mining and machine learning for identifying sweet spots in shale reservoirs. *Expert System with Applications*, 88, 435–447. (in persian)

Tleis, M., Callieris, R., & Roma, R. (2017). Segmenting the organic food market in Lebanon: an application of K-means cluster analysis. *British Food Journal*, 119(7), 1423-1441.

Vadim, K. (2018). Overview of different approaches to solving problems of Data Mining. *Procedia Computer Science*, 123, 234–239.

Xiang, Y., Zuo, Y., & Gaoshan, F. (2024). Design of intelligent detection method for electricity transmission line equipment defect based on data mining algorithm. Thermofluids, 24.