Improve Customer Relationship Management With The Developed RFM Approach In The Communications Industry

1. Maryam Sadeghi [1]
*2. Mohammad Naderi Dehkordi [2]**
3. Behrang Barekatain [3]
4. Naser khani [4]

Recive 2023,11,05 **Accept 2023,12,10**

Abstract

With the competitiveness of the communications industry and the growth of customer expectations, new customer relationship strategies are felt in this industry. Customer Relationship Management is a valuable process in marketing to better understand customers in order to improve profitability and long-term customer relationships. Effective knowledge of customer information can provide services tailored to the functional and behavioral characteristics of each category to improve customer relationship management. Data mining is a powerful tool for companies to extract knowledge from their customers' transaction data. One of the useful applications of data mining is segmentation. In this article, customer transaction data of one of the fixed internet service providers in Isfahan province based on the analysis of the RFM model and the proposed LRFM-Ccr model (length of the customer relationship, recent purchase, number of purchases, paid expenses, and customer consumption records) are selected and are categorized and then clustered using the K Means algorithm. The value of the Davis Boldin index for each number of clusters in the proposed LRFM-Ccr model and its corresponding value in the RFM model have been compared and evaluated. The models are compared with Recall, cluster compactness, accuracy, Number of correct predictions, and correct weight average.

Keywords: RFM Analysis, K Means clustering Algorithm, Customer Relationship Management, Data Mining, Communication Industry

1. *Department of Management, Najafabad Branch, Islamic Azad University, Najafabad, Iran*
2. *Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran. Big Data Research Center, Najafabad Branch, Islamic Azad University, Najafabad, Iran. mndehkordi@iau.ac.ir* ***(Corresponding author)****
3. Faculty of Computer Engineering, Najafabad Branch, Islamic Azad University, Najafabad, Iran. Big Data Research Center, Najafabad Branch, Islamic Azad University, Najafabad, Iran.
4. *Department of Management, Najafabad Branch, Islamic Azad University, Najafabad, Iran*

Introduction

The concept of Customer Relationship Management (CRM) emerged in the early twentieth century. CRM helps an organization acquire knowledge that enables it to better understand customer behavior and relationships, and therefore can retain the customer and improve customer profitability (Sabbeh, 2018; Soltani and Navimipour, 2016). The methods and innovations of data mining and data warehousing in CRM offer the latest approaches for business sectors to adapt to the idea of marketing strategy. Ergo, searching for useful information in the vast amount of data and incorporating the information in practical operations becomes the issue that needs to be tackled promptly, hence contributing to the emerging use of data mining (Sawal et al., 2022). Customer retention strategy in customer relationship management has become increasingly an important issue. Data mining techniques play a vital role in the better implementation of customer relationship management (Haddan et al., 2005). The term data mining refers to the extraction of non-obvious, implicit, previously unknown, and potentially useful information. Data mining has been described as a way to automatically analyze large amounts of data to find patterns and trends that might not be discovered if not used (Hwang et al., 2004). Customer segmentation, which is at the heart of marketing strategy, makes it possible to determine the answers to questions relating to the number of investments to be released, the marketing campaigns to be organized, and the development strategy to be implemented (Doae et al., 2022).

In this article, the use of appropriate data mining techniques for distinguishing and classifying customer relationships with the organization has been investigated. Data mining has been used for various purposes in the communications industry. One of those goals is to segment customers for targeted marketing. From a business perspective, not all customers have the same value. For this reason, to identify key customers, the relevant data is evaluated by organizations so that they can first establish long-term relationships with the customer and secondly, deal with each of them according to the value they create for the organization .Customer segmentation is an important success factor in understanding the behavior of different customer groups and evaluating their value (Yao et al., 2014). It also enables companies to identify and retain valuable customers, which is critical to business success in highly competitive industries (Webster, 1992). Khajvand and Tarokh (2011) consider the RFM model as a market segmentation technique that analyzes customer behavior such as how much a customer has recently bought (delayed), how many times a customer buys (repurchase), and how much money they spend. Slow (purchase volume) is used.

This study presents a CRM strategy that includes customer segmentation and behavior analysis, using the data set of one of the companies providing fixed internet in Isfahan province, which contains about 600,000 customers. RFM model has low clustering quality. On the other hand, the variables of customer relationship Length (L) and customer consumption records (Ccr), in addition to expressing customer performance, there are significant differences between different customers that their inclusion as clustering variables will increase the quality of clustering; Therefore, one of the hypotheses of this research is one of the inefficient quality of RFM model in clustering customers of a company providing fixed internet services and the other is that by adding the characteristics of customer relationship length (L) and customer consumption records (Ccr) to

RFM model and model presentation A new one called LRFM-Ccr will increase the quality of customer segmentation in this industry. To test the research hypotheses, the company's fixed internet customers using their purchase data and based on two models RFM and the proposed model LRFM-Ccr and K Means clustering algorithm, and the quality of the two models in clustering and identifying customer behavior will be evaluated.

This research has two objectives:

(1) It proposes a new RFM model called LRFM-Ccr to better reflect customer behavior in the communications industry.

(2) Its findings on the unique characteristics and behaviors of customers and their consequences for different types of customers, guide communication company managers to improve customer relationships and develop different services and marketing strategies.

The structure of the rest of the paper is as follows. In Section 0 a short review of related work is presented. In Section 0, the research method is presented with a case study of a company providing fixed internet services with managerial results and implications. The research findings are presented in Section 0 and finally, the conclusions are presented in Section 0.

Related work

CRM consists of four dimensions customer identification, customer attraction, retention, and development (Samizadeh, 2008). Zahrotun used CRM to use online customer information to identify the best customer (Zahrotun, 2017). Using the concept of CRM for online shopping, the author identifies potential customers by segmenting customers, which helps us increase the company's profits. Therefore, the Fuzzy C-Means clustering method has been used for customer segmentation and marketing to the customer in an accurate way. Therefore, it helps customers to get special features in more than one category according to their needs, and inappropriate branding strategies.

Jiang and Tuzhilin offer a direct clustering approach that clusters customers not based on calculated statistics but by combining the transaction data of several customers (Jiang and Tuzhilin, 2009). The authors also show that finding an optimal partitioning solution is very difficult. Therefore, Jilin was optimized under different clustering methods. The authors then experimented with customer segments obtained by direct grouping and found that this was better than the statistical method. Learning from large amounts of data is a very challenging issue that most cases of clustering based on methods and other data analysis techniques become more difficult due to the "curse of dimensions" (Bellman, 2015). Therefore, processing large-scale data requires the use of appropriate techniques and methods (Tan et al., 2005).

Cheng and Sun to identify valuable customers in the telecommunications industry, another application of the RFM model called (TFM), which of the three main features to describe users who have accumulated more service time (T), often 3G services (F) and use a large number of invoices per month (M), have used (Cheng and Sun, 2012). Although in other industries, the purchasing freshness variable is important in clustering, in the FMCG industries, this variable has little variability between customers; Because the customers of these industries, due to the nature of the materials purchased, have to buy at short intervals and continuously. On the other hand, the best time for customer clustering and valuation is the end of the fiscal year, which includes regular purchases and customer purchases on special occasions; Because at this time, the clustering error decreases and its accuracy increases (Tsiptsis and Chorianopoulos, 2010).

Several methods have been proposed in the literature to perform customer segmentation, and among them, clustering is the most common method (Wedel and Kamakura, 2012). In addition, the features of the RFM model are used effectively to understand and analyze the characteristics of customer behavior (Fader et al., 2005; Kahan 1998; Newell, 1997) and because of their simplicity and application, it is well-known (Bauer, 1988; Bult and Wansbeek, 1995).

Several versions of RFM models and cluster analysis have been successfully applied together to successfully segment the customer in different application areas. A review of the research background shows that to increase the quality of customer clustering in different industries, researchers have changed the way of measuring clustering variables in the RFM model and have removed or added variables from the model.

RFM model

One of the models in clustering and customer value analysis in customer relationship management is the clustering model, called the RFM model, which was proposed by Hughes (Alizadeh and Karimi, 2018). Peker and Altan (2017) have used the RFM model to classify their specialized customers. This study is very useful in designing marketing strategies for different customers. The basic RFM model relies on three attributes which are recency (R), frequency (F), and monetary (M). Recency is the time interval since the last purchase (e.g., days or months) and gives information about the buying or visiting potential of the customer. If this interval is short, the likelihood of repurchasing or revisiting is high. Frequency is the number of purchases or visits within a certain period, and it is an indication of customer loyalty. The higher the frequency is, the higher the customer loyalty becomes. Monetary is the total amount spent or the average amount spent per visit during a certain period and measures the contribution of the customer to the revenue of a company. The greater the amount spent, the more the customer contributes to the revenue.

K Means Algorithm

The goal of the K Means algorithm is to divide M-points in N-dimensions into K-clusters to minimize the within-cluster sum of squares "Local" optima solution was sought so no movement to a point from one cluster to another will reduce the within-cluster sum of squares the algorithm requires as input a matrix of M-points in N-dimensions and a matrix of K-initial cluster centers in N-dimensions. The number of points in cluster L is denoted by NC(L). D(I, L) is the Euclidean distance between point I and cluster L. The general procedure is to search for a K-partition with the locally optimal within-cluster sum of squares by moving points from one cluster to another (Imani et al., 2022).

LRFM-Ccr Model

Researchers are developing new RFM models by considering additional variables or eliminating some variables depending on the nature of the product or service. Chang and Tsay (2004) created an LRFM model using a new feature, the length of customer relationship (L) in the original RFM model. In the LRFM-Ccr model, customers are classified based on five characteristics: customer relationship length, the novelty of purchase, the number of purchases, the financial value of purchase, and the customer's consumption history, and based on the first letter of each of these five characteristics, the word LRFM-Ccr has been innovated.

Standardization of variables L, R, F, M, Ccr

The scale of each of these features is different, as shown in Table 1. For clustering, the unit of measurement of variables must be the same. In this case, it is necessary to convert the scale and standardize it. To standardize, the maximum-minimum standardization function of Eq. (**Error! Reference source not found.**) is used, based on which the standardized values will be between zero and one. A value of zero will be obtained for the smallest data and a value of one for the largest data.

$$X_i = \frac{X - Y_{min}}{Y_{max} - Y_{min}} \quad i = 1, 2, \dots, n \qquad (1)$$

Y_{min} is the smallest value and Y_{max} is the largest value of the variable and X = L, R, F, M, Ccr will be the main values of the variables, and X_i = L_i, R_i, F_i, M_i, Ccr_i are the standardized values of the main variables.

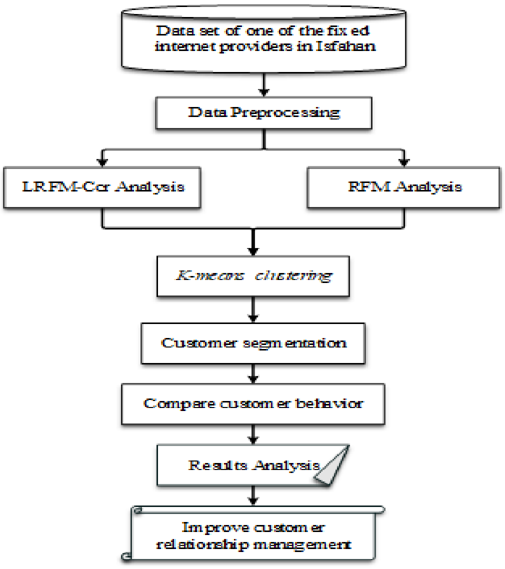
The Proposed Method

The data used in this research is the real data of customers of one of the Internet providers in Isfahan province, Iran. The dataset contains 9453 customer records. RapidMiner software has been used to perform data mining processes. In this study, an efficient combination of the LRFM-Ccr model and clustering is proposed for customer segmentation. One of the most common clustering methods, the K Means algorithm, is applied to customers. In customer databases, the features of the LRFM-Ccr model are hidden in trading data. Therefore, as the first step in customer segmentation, the features of the LRFM-Ccr model are calculated for each customer. Because the actual client data is used for the features of the LRFM-Ccr model, it is assumed that all model variables are of equal weight (i.e. equally important).

Note that the features of the LRFM-Ccr model vary in amplitude, and the difference in scale between these features affects and distorts the results of clustering analysis. So, before clustering, the LRFM-Ccr variables are standardized using,maximum-minimum standardization, based on which the standardized values will be between zero and one. Before using clustering algorithms, finding the optimal number of clusters (k) is an important issue. A common solution to this problem is to use cluster validation indices that provide a score based on the concepts of compactness (ie, members within a cluster are similar) and segregation (i.e., how different members are in different clusters). In this paper, the Davies–Bouldin (DB) index is used to determine the appropriate number of clusters. The DB index estimates compactness based on the distance of all points in the cluster to its center of gravity but defines separation as the distance between centers of gravity. The K Means algorithm is implemented for different numbers of clusters (k) and the results of each Performance are evaluated using the DB index. After deciding the appropriate number of clusters, the clustering results are analyzed with the specified number of clusters. Then, LRFM-Ccr scores are calculated for each customer group, and based on the LRFM-Ccr features, customer segments are indexed, finally, guidelines for improving customer relationship management are proposed in

Fig.1.

***Fig.**1 General research framework*



In this paper, the steps of data mining and data analysis to discover knowledge from them are performed based on the standard CRISP-DM process (Chapman, 2000). This process includes the steps of understanding the business environment, data selection, data preparation, modeling, model evaluation, and results development in

Fig.**2**. The implementation stages of the research include six main steps, which are described below:

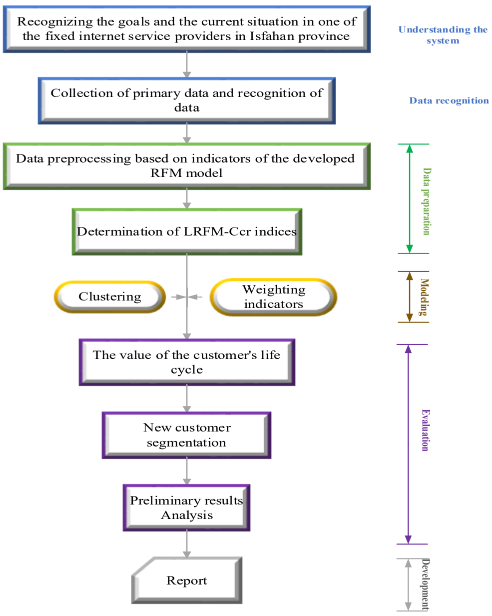
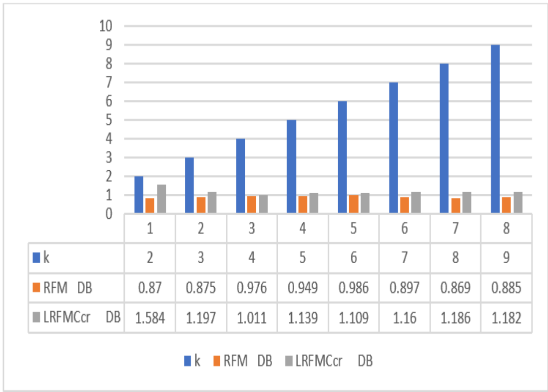


***Fig.2** Steps of Crisp process model in research*



	1	2	3	4	5	6	7	8
k	2	3	4	5	6	7	8	9
RFM DB	0.87	0.875	0.976	0.949	0.986	0.897	0.869	0.885
LRFMCcr DB	1.584	1.197	1.011	1.139	1.109	1.16	1.186	1.182

***Step1.** Understand the business environment;* The company providing fixed internet services is currently providing services in Isfahan province with more than 3000 employees. The company's fixed internet customer base reaches more than 600,000 subscribers. The nature of the type of telecommunication customers is the factor of final consumers and these customers are mainly home, office, commercial, etc. Therefore, the customer clustering approach is very important for recognizing and analyzing the behavior of active customers in this industry. For this reason, this information is used to evaluate and analyze the characteristics of customers and, finally, to provide a relevant marketing mix to formulate a marketing strategy tailored to each segment and achieve the desired results in the field of CRM.

***Step 2.** Select a data set;* First, the sample data related to the basic information and performance data of the customers of this fixed internet service provider company is collected from the CRM database and the database related to the research needs is presented. Sample measurement formulas are different, one of the most widely used methods in determining sample size is

Cochran's formula. Performance data of customers of one of the fixed internet service providers in Isfahan province from 03/20/2020 to 03/20/2021 which includes 9453 customers (more than 0.95% of the statistical population) were collected as a sample.

Step 3. *Data preparation;* At this stage, preparation and preprocessing are done to facilitate the discovery of latent knowledge contained in the data. For this purpose, first incomplete information, data, or invalid and inaccurate values are removed and all data is converted to a format that can be used in Rapidminer data mining software. At this stage, the entire data set is in the ratio of 70 and 30 to educational data. And are divided experimentally.

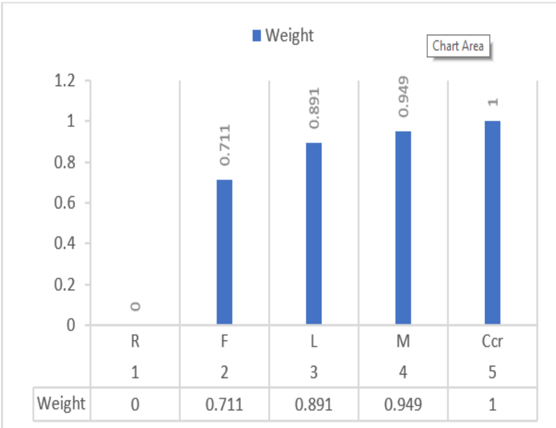
Measurement of clustering variables using Excel software; Per customer, performance variables are measured in two models, RFM and LRFM-Ccr, based on their purchase data. The newness of purchase (R) is equal to the number of days between the last purchase until 03/20/2021, repetition of purchase (F) is equal to the total number of purchases in one year, material purchase value (M) is equal to the total value of Rials of purchase in one year and customer consumption records (Ccr) is equal to the amount of the customer's internet consumption in one year and the length of the customer's relationship with the company (L) indicates the time that a customer has started his relationship with the company (Table 1).

Table 1 Method of measuring the variables L, R, F, M, Ccr

Symbol	Variable	Definition	the unit
L	Length of customer relationship	Number of days from the first purchase to the last purchase	Number of days
R	Recently bought	The number of days since the last purchase until now	Number of days
F	Number of repeat purchases	Number of purchases during the period	times
M	Value for Money	Money volume purchased during the period	Rial
Ccr	Customer consumption records	The amount of fixed internet consumed by the customer during the period	the amount of

Step 4. *Modeling, and selecting the number of clusters (k);* There are several types of useful indexes for determining k as the number of clusters. In this study, DB index was used to find the optimal number of clusters. The purpose of the DB index is to identify cluster sets that have small intra-cluster distances and large inter-cluster distances. The design process for selecting the number of clusters (k) and the value of the DB index was performed using Rapidminer software. The value of DB and the comparison graph of DB index for each number of clusters, for the values of k equal to 2,3,4, ..., 9, are shown in Table 2 and **Error! Reference source not found.**.

Fig.3 Comparison diagram of DB index for each number of clusters (k)



Clustering customers with K Means algorithm; At this stage, for a range of a number of clusters, using the K Means algorithm and considering the clustering variables based on two RFM and LRFM-Ccr models, the company's customers are clustered according to the minimum value of the DB index in Table 2 (i.e. k = 8 for the RFM model and k = 4 for the LRFM-Ccr model). The results of experimental data clustering are shown in Table 3 and Table 4.

Table 2 DB value for each number of clusters (k)

k	RFM DB	LRFMCcr DB
2	0.87	1.584
3	0.875	1.197
4	0.976	1.011
5	0.949	1.139
6	0.986	1.109
7	0.897	1.16
8	0.869	1.186
9	0.885	1.182

Table 3 Results of K Means clustering in RFM model and standard values of variables

Attribute	cluster_0	cluster_1	cluster_2	cluster_3	cluster_4	cluster_5	cluster_6	cluster_7
F	0.332	-0.844	0.771	1.105	-0.409	-0.736	-0.662	4.017
M	-0.352	-0.458	2.876	0.743	0.799	-0.255	-0.650	2.144
R	-0.567	3.027	-0.381	-0.543	-0.126	1.338	-0.371	-0.557

Table 4 Results of K Means clustering in LRFM-Ccr model and standard values of variables

Attribute	cluster_0	cluster_1	cluster_2	cluster_3
F	-0.108	-0.772	-0.048	1.358
Ccr	-0.041	-0.044	-0.038	-0.020
L	-0.346	-0.193	1.982	-0.238
M	-0.320	-0.246	-0.091	1.481
R	-0.450	1.759	-0.254	-0.498

LRFMCcr weight determination by Rapidminer software; In this paper, Rapidminer software is used to calculate the weight of LRFMCcr according to the opinions of decision-makers. This was done in 2 steps. In the first step, three decision-makers are selected from three different sales management layers for pairwise comparisons. They include a senior manager, a middle manager, and an operational manager. In the second step, the incompatibility index was calculated and evaluated for each decision-maker. Finally, the weight of LRFM-Ccr is determined by calculating the specific value using the Calculate Weighths operator according to **Table 5** and the weight diagram of the LRFM-Ccr variables are shown in Error! Reference s ource not found.**.**

Table 5 Displays the weight of the values of the LRFM-Ccr model variabl

Row NO.	Weight	Attribute
1	0	R
2	0.711	F
3	0.891	L
4	0.949	M
5	1	Ccr

Fig.4 Graph showing the weight of LRFM-Ccr variables

PerformanceVector:
accuracy: 96.24% +/- 1.02% (micro average: 96.24%)
ConfusionMatrix:

True:	cluster_1	cluster_0	cluster_3	cluster_2
cluster_1:	530	3	2	5
cluster_0:	5 1441	40	6	
cluster_3:	2	12	355	9
cluster_2:	6	4	10	335

kappa: 0.941 +/- 0.016 (micro average: 0.941)
weighted_mean_recall: 94.47% +/- 1.73% (micro average: 94.47%), weights: 1, 1, 1, 1
weighted_mean_precision: 95.82% +/- 1.37% (micro average: 95.75%), weights: 1, 1, 1, 1
spearman_rho: 0.950 +/- 0.020 (micro average: 9.501)
correlation: 0.948 +/- 0.020 (micro average: 0.948)

Step 5- *Evaluation;* Ranking the customer life cycle value (CLV) of departments based on LRFM-Ccr values and related weights can help managers identify more important departments. They can plan for customer retention based on this ranking. To determine the LRFM-Ccr rating for sections, the values of L, R, F, M, and Ccr were first standardized for centers of gravity. The LRFM-Ccr score for each section is then calculated as described in Eq. (2) and the results are displayed in Table 6 and Table 6 **Value of customer life cycle and ranking of clusters in RFM model**

Table 7.

$$C^j = w_L C_L^j + w_R C_R^j + w_F C_F^j + w_M C_M^j + w_{Ccr} C_{Ccr}^j \quad (2)$$

Table 6 Value of customer life cycle and ranking of clusters in RFM model

Cluster	CLV	Percentage of members of each cluster	Number of members per group	Rank
C0	-0.097996	0.295	816	5
C1	-1.034726	0.049	136	7
C2	3.277505	0.0347	96	2
C3	0.319034	0.129	357	4
C4	0.467452	0.092	256	3
C5	-0.4233	0.156	431	6
C6	-1.087532	0.227	630	8
C7	4.890743	0.015	43	1

Table 7 Customer lifetime value and cluster ranking in the LRFMCcr model

Cluster	CLV	Percentage of members of each cluster	Number of members per group	Rank
C0	-1.0432282	0.528	1460	4
C1	-0.998309	0.196	543	3
C2	1.607475	0.128	355	2
C3	2.138949	0.147	407	1

In this equation, C^j is equal to the rank of LRFM-Ccr in cluster j and C_L^j, C_R^j, C_F^j, C_M^j, and C_{Ccr}^j are the standardized values of the variables L, R, F, M, Ccr in the cluster *j*, and w_L, w_R, w_F, w_M, and w_{Ccr} are the weights of the variables L, R, F, M, Ccr in the j cluster.

For cluster analysis, R, F, L, Ccr, and M parameters should be classified into four categories (low, medium, high, and very high). These categories are determined with the coordination of the company expert, as shown in Table 8. By comparing the results of the LRFM-Ccr parameter values in each cluster with the classified values, shown in Table 6 **Value of customer life cycle and ranking of clusters in RFM model**

Table 7 and Table 8, the category of each parameter is determined to determine the CLV rank for each cluster in Table 9. Also, the predictive performance of the LRFM-Ccr model can be seen using the output results of the cross-validation operator and the Performance sub-process according to the clustering design process in Rapidminer software, in

Table *10*, Table 11 and Fig.5.

Step 6. Development of results; The findings in Table 9 provide unique guidance to company executives on improving customer relationships and developing different marketing services and strategies, regarding the unique characteristics and behaviors of customers and their consequences for different types of customers.

Table 8 LRFM-Ccr Categories

	Low	Medium	High	Very High
L	0 269	270 774	775 1053	1054 3511
R	0 9	10 30	31 120	121 365
F	1 3	4 8	9 14	15 131
M	500 1427500	1428332 2393333	2395000 3503166	3505130 2936332
Ccr	168 245760	246016 450560	450816 777144	777145 1228800000

Table 9 LRFM-Ccr classification and CLV Ranking for each cluster

Cluster	LRFM-Ccr Analysis					
	L	R	F	M	Ccr	CLV Rank
C0	Low	Medium	Medium	Low	Medium	4
C1	High	Very High	Low	Medium	Low	3
C2	Very High	High	High	High	High	2
C3	Medium	Low	Very High	Very High	Very High	1

Table 10 Results of clustering experimental data of RFM model and k = 8 in K Means

accuracy: 94.68% +/- 1.08% (micro average: 94.68%)

	true cluster_1	true cluster_6	true cluster_0	true cluster_3	true cluster_2	true cluster_5	true cluster_4	true cluster_7	class precision
pred. cluster_1	136	0	0	0	0	2	0	0	98.55%
pred. cluster_6	0	609	11	0	0	1	4	0	97.44%
pred. cluster_0	0	17	780	37	0	0	7	0	92.75%
pred. cluster_3	0	0	19	313	4	0	4	0	92.06%
pred. cluster_2	0	0	0	2	89	1	0	1	95.70%
pred. cluster_5	0	3	0	2	0	420	12	0	96.11%
pred. cluster_4	0	1	6	0	2	7	229	0	93.47%
pred. cluster_7	0	0	0	3	1	0	0	42	91.30%
class recall	100.00%	96.67%	95.59%	87.68%	92.71%	97.45%	89.45%	97.67%	

Table 11 Clustering results of experimental data of LRFMCcr model and k = 4 in K Means

accuracy: 96.24% +/- 1.02% (micro average: 96.24%)

	true cluster_1	true cluster_0	true cluster_3	true cluster_2	class precision
pred. cluster_1	530	3	2	5	98.15%
pred. cluster_0	5	1441	40	6	96.58%
pred. cluster_3	2	12	355	9	93.92%
pred. cluster_2	6	4	10	335	94.37%
class recall	97.61%	98.70%	87.22%	94.37%	

Table 12 Results of the final classification taken from Fig. 5

Metrics	Weighted Mean Accuracy	Weighted Mean Recall	Weighted Mean Precision	Kappa
RFM	94.68	94.61	94.93	0.934
LRFM-Ccr	96.24	94.47	95.82	0.941

Findings

The results of the final classification according to Table 12 show that the accuracy of the LRFM-Ccr model is 96.24%, 1.56% higher than the RFM model, and due to the increase in kappa in the proposed model (0.941), the prediction accuracy is 0.007% Increased (kappa = 0, all predictions are majority class. Kappa = 1, all predictions are correct. Kappa = −1, all predictions are wrong) and also the weighted average of the proposed model (95.82) compared to The RFM model has increased by 0.89%.

Discussion and Conclusion

An LRFM-Ccr model was presented for the segmentation of fixed internet services. 4 clusters were obtained using the K Means algorithm with optimal k according to the Davies-Bouldin index on LRFM-Ccr changes. The obtained clusters were divided by weighting LRFM-Ccr and the weight of each amount was calculated using Rapidmin software. In addition, each cluster is analyzed and performed. The results show an increase in the quality of segmentation of customers in this industry with the new model. The results show that the proposed model is equal to the RFM model in repetition, in more compact cluster compaction, with increasing accuracy, with increasing prediction accuracy, and with increasing average weight. The results of this research according to Table 8 and Table 9 can help marketing managers gain in-depth insights into customers and provide guidelines for improving customer relationships. The limitation of this work is the collection of reliable data from random customers (only about 10,000 customers). Due to following direct methods and limited human resources, this process took a long time (about 7 months). Another limitation is that this has been applied in Isfahan province, which may be different from other provinces or even other communities, so the instructive features in this study can be more or less important in other communities. Future work includes studying customer performance in other features, which may help provide better suggestions for improving customer relationship management.

Acknowledgments

The authors thank the data provider for their extensive support.

References

- Abdolmajid Imani, Meysam Abbasi, Farahnaz Ahang, Hassan Ghaffari1 and Mohamad Mehdi. 2022. “Customer Segmentation to Identify Key Customers Based on RFM Model by Using Data Mining Techniques.” International Journal of Research in Industrial Engineering.
- Ainul Balqis Sawal, Muneer Ahmad, Mitra Anusri Muralitharan, Vinoshini Loganathan and N. Z. Jhanjhi. 2022. “Machine Intelligence in Customer Relationship Management in Small and Large,Companies.”. https://doi.org/10.4018/978-1-7998-9201-4.ch007.
- Ali Alizadeh Zoeram , AhmadReza Karimi Mazidi. 2018. “A New Approach for Customer Clustering by Integrating the LRFM Model and Fuzzy Inference System.” Iranian Journal of Management Studies (IJMS). Spring 2018. Vol. 11, No. 2, pp. 351-378.
- Bauer, C.L. 1988. “A direct mail customer purchase model, Journal of Direct Marketing.” Wiley Online Library, Vol. 2 No. 3, pp. 16–24.
- Bellman RE. 2015. “Adaptive control processes: a guided tour.” vol. 2045. Princeton: Princeton University Press.
- Bult, J.R. and Wansbeek, T. 1995. “Optimal selection for direct mail, Marketing Science” INFORMS, Vol. 14 No. 4, pp. 378–394.
- Chang, H. H., & Tsay, S. F. 2004. “Integrating of SOM and K-means in data mining clustering: An empirical study of CRM and profitability evaluation.” Journal of Information Management, 11(4), 161-203.
- Cheng L-C, Sun L-M. 2012. “Exploring consumer adoption of new services by analyzing the behavior of 3G subscribers: an empirical case study.” Elect Comm Res Appl.;11 (2):89–100.
- Doae Mensouri, Abdellah Azmani, Monir Azmani. 2022. “K-Means Customers Clustering by their RFMT and Score Satisfaction Analysis.” International Journal of Advanced Computer Science and Applications, Vol. 13, No.6.
- Fader, P.S., Hardie, B.G.S. and Lee, K.L. 2005. “RFM and CLV: Using iso-value curves for customer base analysis.” Journal of Marketing Research, American Marketing Association, Vol.42 No. 4, pp. 415–430.
- Hadden, Tiwari, R.Roy, D. Ruta. 2005. “Computer-assisted customer churn management: State-of-the-art and future trends.” Computers &Operations Research.
- Hwang, Jung & Suh. 2004. “An LTV model and customer segmentation based on customer value: a case study on the wireless telecommunication industry.” Expert Systems with Applications.
- Jiang, T., Tuzhilin, A. 2009. “Improving personalization solutions through optimal segmentation of customer bases. ” IEEE Trans. Knowledge Data Eng. 21(3), 305–320. https://doi.org/10.1109/TKDE.2008.163N.
- Kahan, R. 1998. “Using database marketing techniques to enhance your one-to-one marketing initiatives.” Journal of Consumer Marketing, MCB UP Ltd, Vol. 15 No. 5, pp. 491–493.
- M. Khajvand and Tarokh.: 2011. “Estimating customer future value of different customer segments based on the adapted RFM model in a retail banking context.” Procedia Computer Science 3 :1327-1332.
- Newell, F. 1997. “The New Rules of Marketing: How to Use One-To-One Relationship Marketing to Be the Leader in Your Industry.” McGraw-Hill, Inc., New York.
- P. Chapman, J. Clinton, R. Kerber, T. Khabaza, T. Reinartz, C. Shearer, R. Wirth. 2000. “CRISP-DM 1.0: Step-by-step data mining guide.”
- Reza Samizadeh. 2008. “Data mining and customer relationship management. ” Rooshd of Andishe Publishing.
- Sabbeh, S.F. 2018. “Machine-learning techniques for customer retention: a

comparative study.” Int. J. Adv. Comput. Sci. Appl. 9 (2), 273–281.

- Serhat Peker and Altan Koçyiğit. 2017. “LRFMP model for customer segmentation in the grocery retail industry: a case study.” Journal of Marketing Intelligence & Planning.https://doi.org/10.1108/MIP-11-2016-0210.
- Soltani, Z., Navimipour, N.J. 2016. “Customer relationship management mechanisms: a systematic review of the state of the art literature and recommendations for future research.” Comput. Hum. Behav. 61, 667–688.
- Tan P-N, Steinbach M, Kumar V. 2005. “Introduction to data mining.” Boston: Addison-Wesley Longman Publishing Co.
- Tsiptsis, K. & Chorianopoulos A. 2010. “Data Mining Techniques in CRM: Inside Customer Segmentation.” WILEY. United Kingdom.
- Webster Jr, F.E. 1992. “The changing role of marketing in the corporation.” The Journal of Marketing, JSTOR, pp. 1–17.
- Wedel, M. and Kamakura, W.A. 2012. “Market Segmentation: Conceptual and Methodological.”.
- Yao, Z., Sarlin, P., Eklund, T. and Back, B. 2014. “Combining visual customer segmentation and response modeling.” Neural Computing and Applications, Springer, Vol. 25 No. 1, pp. 123–134.
- Zahrotun, L. 2017. “Implementation of data mining technique for customer relationship management (CRM) on online shop tokodiapers.com with fuzzy c-means clustering. ” In: 2017 2nd International Conferences on Information Technology, Information Systems and Electrical Engineering (ICITISEE).