Dynamic customer segmentation: a case study using the modified dynamic fuzzy *c*-means clustering algorithm

M. Sivaguru[1] (ID)

Abstract

Dynamic customer segmentation (DCS) is a useful tool for managers to adjust their marketing strategies from time to time. However, no study in the literature has attempted to develop a DCS framework until now. To fill the research gap, a DCS framework is proposed. To improve the effectiveness of the proposed framework, the existing dynamic fuzzy *c*-means clustering (dFCM) algorithm is modified owing to certain limitations found in it. Extensive experiments were conducted using the retail supermarket dataset to assess the performance of the modified dFCM (MdFCM) algorithm. Experimental results prove that the MdFCM algorithm performs better than the existing algorithm. The experimental results are validated by fuzzy clustering evaluation measures such as Xie–Beni index (XBI), partition coefficient (PC), modified partition coefficient (MPC), partition entropy (PE), and fuzzy silhouette index (FSI). A statistical significance test, MANOVA Pillai's statistics, is carried out to prove that clusters obtained from the MdFCM algorithm are significant. Finally, a case study on a retail supermarket has been conducted using the proposed DCS framework. The study has shown that the proposed DCS framework extracts useful information for managers to support strategic decision-making.

Keywords Dynamic customer segmentation · Dynamic fuzzy *c*-means · RFM analysis · Retail supermarket

1 Introduction

Customer segmentation is an important activity of customer relationship management (CRM) to divide a broad customer base into manageable sub-groups of customers. The segmentation concept is introduced by Smith (1956). The concept assumes that every customer's needs and expectations are different; therefore, they require different marketing strategies to fulfil their needs (Wedel and Kamakura 2000). In recent times, owing to the proliferation of database technologies in the retail industry, wherein owners of retail supermarkets tend to maintain large customer transaction databases, customer segmentation has attracted significant academic interest (Chan et al. 2011). In particular, customer segmentation in the retail supermarket industry has significantly impacted strategic decision-making (Teichert et al. 2008). Customer segmentation

helps managers understand the customer needs and requirements for formulating effective marketing strategies (Hiziroglu 2013). Managers or owners of retail supermarkets are always interested in segmenting their customers for various reasons. For example, when there are many customers in a retail supermarket, managers have to divide their large customer base based on their buying pattern, i.e., low-value or high-value customers. This process helps the organization (1) to increase customer retention and loyalty, (2) customization of products and services (3) increase profitability (Kim et al. 2006). In the same vein, customers who are supermarket buyers will get personalized services such as special discounts. Segmenting the customer base gives critical information on customers' buying patterns and experiences, which can be useful for product design, product recommendation, and better customer services. Furthermore, efficient customer segmentation gives marketers an edge over competitors in sustaining their best customers, and improving their potential customers (Bachtiar 2019).

Customer segmentation is a process of similar grouping customers based on common characteristics such as similar

✉ M. Sivaguru
   sivaguru300@gmail.com

1 GITAM School of Business, GITAM University,
   Bengaluru 561203, India

buying patterns found among the customers. It is typically used to identify highly profitable customers or customers who have the potential to become profitable customers (Han et al. 2012). It extracts useful information such as the percentage of profitable customers and their contribution to total sales, i.e., the Pareto principle (You et al. 2015). Consequently, customer segmentation helps formulate marketing strategies for each segment identified (Zeybek 2018). Not only is it difficult to target each of the customers individually, but also it is challenging for companies to manage a large customer base. Therefore, it is necessary to develop new frameworks that managers can use as a tool in customer segmentation. In line with this, one of the methods introduced in the literature to do customer segmentation is the recency, frequency, and monetary (RFM) analysis method (Hughes 1994). The notion behind the analysis is that customers who bought recently, frequently, with high monetary value would buy again. Segments formed through RFM analysis can help formulate marketing strategies to achieve organizational goals (Fan and Zhang 2009).

Segmenting the customers mainly depends on clustering algorithms to find natural groupings among the customers. Many potential customer segmentation applications are not feasible because of the poor performance of clustering algorithms (Arunachalam and Kumar 2018). Customer transaction-related data are typically clustered using clustering algorithms. These algorithms can be classified into two different categories. One category of algorithms deals with well-separated hard clustering problems wherein there is no overlap between the clusters, e.g., k-means clustering algorithm (Macqueen 1967). Another category of algorithms deals with overlapping soft clustering algorithms wherein data points are assigned to multiple clusters, e.g., fuzzy c-means clustering algorithm (Bezdek et al. 1984). (Turksen 1986) proposed the theory of interval-valued fuzzy sets based on normal forms. Some applications of interval-valued fuzzy sets have been presented (Chen 1997; Chen et al. 1997; Chen and Hsiao 2000).

The fuzzy c-means clustering algorithm is one of the widely used algorithms to achieve the customer segmentation objective (Hiziroglu 2013). One of the drawbacks of these algorithms is that they are static, i.e., clusters obtained from these algorithms cannot be updated by new information. Hence, segmentation done using these algorithms becomes outdated after some time. Also, it is affected by noise and outliers (Askari 2021). For example, a retail supermarket that tries to segment its broad customer base cannot rely upon the static customer segmentation results; instead, they require a customer segmentation system that can continuously update its system with new information to study changes in the customer segments. Dynamic clustering algorithms were developed to update

the clusters with new information (Seret et al. 2014). These algorithms operate across different time scales to update customer segments with new information. One of the algorithms capable of updating the customer segments with new information is the dFCM algorithm (Fathabadi 2016).

The dFCM algorithm was used for various applications ranging from evaluating stock performance to reconfiguring the power distribution networks. However, the algorithm has shortcomings that prevent its successful application to the DCS problem. First, the algorithm is not used for the DCS problem until now in the literature. Most importantly, their performance in identifying the changes in the cluster structure effectively in the context of DCS is never investigated before in the literature. Second, the algorithm suffers from the user-defined parameter 'Y_{thr}' used to identify the changes in the cluster structure, i.e., movement of clusters (c), creation of new clusters $(c + 1)$, and elimination of clusters $(c - 1)$. It is to be observed that the user fixes it, and there is no method available in the literature to calculate the optimal value of this parameter. Poor choice of this parameter may result in lousy clustering results.

This study aims to overcome the shortcomings mentioned above for doing the DCS effectively. The major contributions of the study are listed as follows.

1. To overcome the first shortcoming of using the dFCM algorithm for the DCS problem, a novel DCS framework based on the dFCM algorithm is proposed in this study. Managers can use this framework as a tool for DCS. The framework is organized into three consecutive phases. In phase-I, a retail supermarket dataset of one year is clustered in cycles using dFCM clustering algorithm. In phase II, clusters obtained from phase-I are classified into different customer segments based on customers' buying patterns. In phase III, customer targeting strategies are devised for each segment based on the changes observed in the customer segments over time.

2. To overcome the second shortcoming of user-defined parameter 'Y_{thr}' used to identify the changes in the cluster structure; the dFCM algorithm is modified using the distance calculated between the new data points and the existing clusters. And it is called the MdFCM algorithm. The modified algorithm can effectively remove the burden of the user-defined parameter.

Extensive experiments were conducted using a retail supermarket dataset of one year with eleven cycles of new updates to assess the performance of the MdFCM algorithm. The performance of the MdFCM algorithm is validated by fuzzy clustering evaluation measures such as Xie-Beni index (XBI), partition coefficient (PC), modified

partition coefficient (MPC), partition entropy (PE), and fuzzy silhouette index (FSI). A statistical significance test using MANOVA Pillai's statistics is carried out to check whether clusters obtained from the MdFCM algorithm are significant or not. Finally, a case study on a retail supermarket has been conducted to demonstrate how our framework can help managers change marketing strategies by considering the changes detected in the customer segments over time.

The remaining part of this paper is organized as follows. Section 2 discusses the related works of literature. Section 3 gives a brief review of existing algorithms. Section 4 introduces our proposed methods. Section 5 talks about the experiments conducted. Section 6 discusses the results obtained from the experiments. Section 7 discusses the case study conducted. Finally, Sect. 7 concludes the paper with remarks.

2 Literature review

The section is divided into two subsections. The first section discusses the DCS and its related concepts. The second section discusses the clustering algorithms used in the context of DCS.

2.1 Background of DCS

DCS is generally known as updating the segments with new information. The process helps in the decision-making process by deriving intelligence that is up to date (Peters et al. 2012). This concept came into being to overcome the limitations attached to static customer segmentation (SCS). One limitation of SCS is that it cannot be updated with new information. Moreover, the knowledge derived from SCS becomes obsolete later (Cachon 2012). To overcome these limitations, a study was conducted that focused on identifying the changes happening in the segments over time. It was identified through the cluster's appearance, disappearance, merging, moving, shrinking, and growing phenomenon. Eventually, it was proved by studying the customer's attitudes over time that it is indeed dynamic. The dynamic pricing model is one of the most widely used applications in this domain (Zhang et al. 2017). However, some studies deal with studying customer profitability over time. They start as low-profit customers but eventually develop into high-profit ones (Rust et al. 2011). A dynamic CRM model is developed using data mining algorithms to solve critical managerial problems. The problem of studying the growth of a loyal customer base is solved in this study (Ha et al. 2002). A model has been proposed for profiling the customer's preferences dynamically. The knowledge derived from the model is used for customizing

real-time financial information products to customers. Consumer preference for a retail store is more dynamic, and studying it helps formulate customer loyalty programs (Lim and Lee 2015).

We have to apply the clustering concept to find the natural groupings in the dataset. The following subsection discusses the clustering algorithms used for DCS problems.

2.2 Clustering in the field of DCS

In our review, we found that dynamic clustering algorithms have been used rarely to achieve the objective of DCS. The reason there is a shortage is the lack of attention given by the researchers. But we have literature that deals with applying fuzzy sets to inventory problems (Garai et al. 2019). Even though there is a growing need for developing dynamic clustering algorithms to fulfil the application requirements, there is a literature shortage addressing this gap. For example, a study conducted by (Hu and Zhao 2015) used a rough k-means clustering algorithm to cluster the electricity customers over time. The segment size and change in the roughness of clusters were used to study electricity customer dynamics. Insurance customers were dynamically clustered using the dynamic rough clustering algorithms to demonstrate the utilities of using it for DCS (Peters and Weber 2012). Based on the fuzzy c-means (FCM) clustering algorithm, a fuzzy recommender system was developed for predicting the user's behavior dynamically (Nadi et al. 2010). Using the latent Dirichlet model (LDA), the temporal behavior sequence of mobile customers was studied (Wilson et al. 2018). A summary of the literature is given in Table 1.

After reviewing the current research, we identified the literature gaps in the following aspects. To the best of author's knowledge, no study has combined the dFCM algorithm with the DCS problem. To fill this research gap, we have proposed a novel DCS framework. As part of that process, we have proposed the MdFCM algorithm.

3 Existing algorithms

In this section, first, the FCM algorithm is introduced. In the second section, the dFCM algorithm is introduced. For an easy understanding of the symbols and notations used in this study, Table 2 is given below.

3.1 FCM algorithm

The algorithm is introduced by Bezdek et al. (Bezdek et al. 1984). The algorithm proceeds as follows.

Table 1 Summary of literature review

The objective of the study	Algorithm	References
Dynamically deriving behaviorally persistent segments	k-means	Reutterer et al. (2006)
Personalizing the products for customers based on the available new information about the customers	k-means	Bernstein et al. (2016)
Detecting temporal changes in customer behavior	FCM	Bose and Chen (2014)
Studying the customers switching patterns between the segments	Self-organizing map	Yao (2013)
Detecting how customers move between clusters over time	FCM	Bose and Chen (2015)
Understanding the migration pattern of customers	k-means	Lingras et al. (2005)
Analyzing the purchase behavior of customers	k-means	Anitha and Patil (2020)

Table 2 Symbols and notations

Symbols and notations	Description
E_z	Objective function of minimizing within the sum of squared error
c	The number of clusters
n	The number of observations
X_j	The feature vector of observations $j, j = 1, \ldots, n$
V_i	Cluster centers $i, i = 1, \ldots, c$
$\overline{X}$	Feature vector of the mean of a dataset
μ_{ij}	Degree of membership of observations j to cluster $i, i = 1, \ldots, c, j = 1, \ldots, n$
z	Fuzzifier to be fixed by the user
d^2	Euclidean distance between observation j and cluster i
X_k	New observations, where $k = n + 1, \ldots \ldots, n + m$
m	The total number of new observations
$\widehat{\mu}_{ik}$	Membership of new observation k to clusters i
$\widehat{d}_{ik}$	Distance between new observation k to clusters i
Y_{thr}	Maximum membership threshold
E_{FCM}	Error threshold used to check the convergence of algorithms
$c + 1$	Creation of a new cluster
$c - 1$	Elimination of cluster
λ_i	Eigenvalues where $i = 1 \ldots n$
H	Hypothesis sums of squares and cross products matrix
E	Error sums of squares and cross products matrix

The objective function of the FCM algorithm E_z is to be minimized with respect to fuzzy membership μ_{ij} and cluster center V_i.

$$E_z = \sum_{i=1}^{c} \sum_{j=1}^{n} (\mu_{ij})^z d^2(X_j, \ V_i). \tag{1}$$

Step 1: Initialize memberships μ_{ij} of X_j belonging to cluster i such that

$$\sum_{i=1}^{c} \mu_{ij} = 1. \tag{2}$$

Step 2: Compute the cluster center V_i for $i = 1, 2, \ldots, c$ using

$$V_i = \frac{\sum_{j=1}^{n} (\mu_{ij})^z X_j}{\sum_{j=1}^{n} (\mu_{ij})^z}. \tag{3}$$

Step 3: Update the fuzzy membership μ_{ij} using

$$\mu_{ij} = \frac{\left(\frac{1}{d^2(X_j, V_i)}\right)^{\frac{1}{(z-1)}}}{\sum_{i=1}^{c} \left(\frac{1}{d^2(X_j, V_i)}\right)^{\frac{1}{(z-1)}}}. \tag{4}$$

Step 4: Repeat steps 2 and 3 until the value E_z is no longer decreasing.

3.2 dFCM algorithm

The dFCM algorithm is introduced by Fathabadi (Fathabadi 2016). It proceeds as follows.

1. For the incoming new data updates, the membership threshold (Y_{thr}) and the convergence criterion (E_{FCM}) are fixed as a first step.
2. The 'c' cluster centers are uniformly located in the input space, and the memberships of the initial data are calculated using Eq. (4).
3. Now, new data updates 'k' is entered in the algorithm, and its memberships are calculated using Eq. (4). If the maximum membership value is greater than or equal to the membership threshold (Y_{thr}), it indicates that it belongs to the existing 'c' clusters. So, cluster the new data updates into the existing clusters by following the iterative process until it satisfies the convergence criterion (E_{FCM}).
4. If it is less than the membership threshold, examine is there any better choice available than the existing 'c' clusters by using the XB clustering validity index (see Eq. 5), i.e., check the validity index of '$c-1$' and '$c+1$' clusters.

$$XB = \frac{\sum_{i=1}^{c} \sum_{k=n+1}^{m} \mu_{ik}^2 \left\| V_i - X_k^2 \right\|}{n \underset{i,p}{\min} \left\| V_i - V_p^2 \right\|}. \tag{5}$$

5. If '$c-1$' or '$c+1$' has a better validity index, '$c-1$' or '$c+1$' clusters are created using new data updates. The clustering process ends when there are no new data updates.

Evaluating cluster validity

Assume that V_u and V_i are the new and existing cluster center vectors, respectively. Now, a condition is defined in Eq. (6):

$$\left\| V_u - V_i \right\| = \sqrt{\sum_{i=1}^{c} (V_u - V_i)^2} > Y_{thr}. \tag{6}$$

Evaluating the cluster validity is necessary if the above condition is satisfied. It is introduced to avoid unnecessary checking of clustering validity. If the new data is present nearer to existing clusters, then it is not checked whether the '$c+1$' clusters are better or not.

The problem with this algorithm is it suffers from the user-defined Y_{thr} threshold. When the algorithm is applied to the DCS problem, the user using it will be confronted with the problem of finding the optimal value of the threshold. No method in the literature would help the user to fix the threshold. Because of that, the user may not have confidence in his clustering results. To address the problem, we have proposed the MdFCM algorithm that would solve the problem. The following section discusses the DCS framework and MdFCM algorithm.

4 Proposed methods

The main objective of this study is to propose a novel DCS framework. As part of that process, we propose an MdFCM algorithm to improve the effectiveness.

4.1 DCS framework

The field of business analytics has become an important aspect of decision support systems to improve the decision-making purpose constantly. While SCS is well accepted in the literature, the typical approach does not consider its dynamic changes, i.e., it cannot update the customer segments over time with new information (Gür Ali and Aritürk 2014). On the other hand, DCS would allow the customer segments to be updated with the latest information. From the point of view of managers of the retail supermarket, managers need insights into the growth of their customer segments to implement appropriate marketing strategies (Bansal et al. 2005). Further, studying customer segments' growing and shrinking behavior may enable managers to contain the growth of undesirable segments (Allaway et al. 2014). Finally, DCS is necessary to adjust organizations' marketing strategies from time to time (Khan et al. 2009).

A dynamic clustering-based DCS framework is proposed to explore the buying pattern of customers from several aspects. The proposed framework, as given in Fig. 1, comprises three consecutive phases. The first phase involves clustering the new data updates using the MdFCM algorithm to find the changes in the cluster structure. In the second phase, clusters obtained from the first phase are classified into different segments based on the RFM pattern of clusters. In the last phase, marketing stages are devised for each segment based on the changes identified in the segments. The following elaborates on the three phases in detail.

4.1.1 Phase-I using the proposed MdFCM algorithm

In this section, we propose an MdFCM algorithm to overcome the shortcomings of user-defined parameter 'Y_{thr}' of the dFCM algorithm. The algorithm is modified by incorporating the condition given in Eq. (7). The condition mainly functions around calculating the distance between the new data updates and existing clusters.

$$d_{ik} > \min\{d(V_i, V_j)\} \tag{7}$$

Fig.1 DCS framework
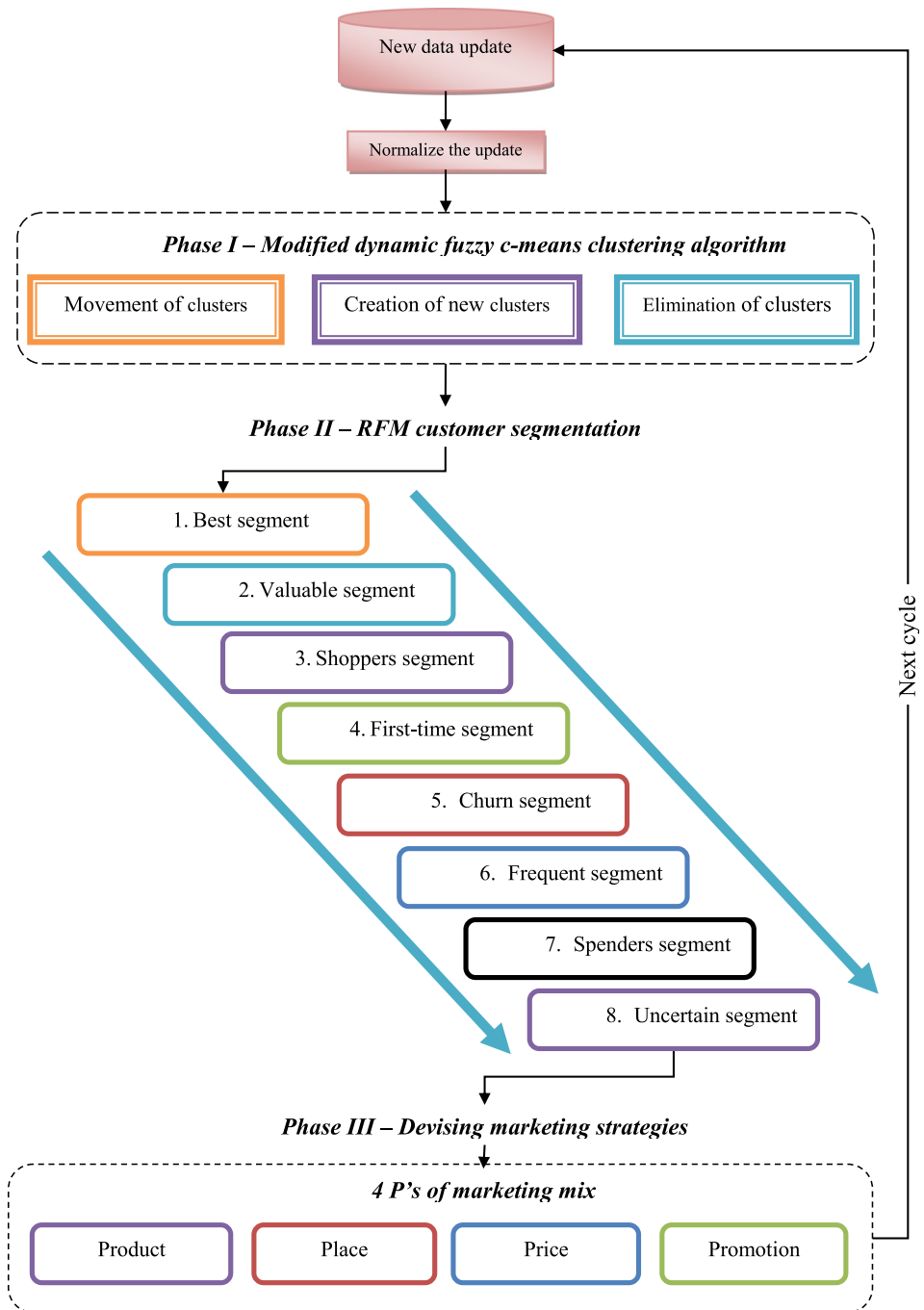
If the calculated distance of new data updates 'd_{ik}' is greater than the minimum distance calculated between the existing clusters; it indicates that the new data updates are somewhere away from the existing clusters; therefore, it requires either creating new clusters or eliminating clusters. If it is less than that, movement of clusters is required. It is to be highlighted that the condition defined in Eq. (7) can effectively solve the problem attached to the user-defined parameter 'Y_{thr}' of dFCM algorithm, i.e., whether to create new clusters $(c+1)$ or eliminate clusters $(c-1)$. The steps involved in the MdFCM algorithm are as follows.

1. Calculate the distance between the existing cluster centers
2. Choose the minimum distance among the distances calculated in step 1
3. Now calculate the distance of new data points to the existing clusters
4. Check the distance of new data points with the distance calculated in step 2 (see Eq. 7)

5. If it is greater than that, go for the creation of new clusters $(c + 1)$ or elimination of clusters $(c - 1)$
6. If '$c - 1$' or '$c + 1$' has a better validity index, '$c - 1$' or '$c + 1$' clusters are created using new data updates.
7. Otherwise, go for movement of clusters (c)
8. The clustering process ends when there are no new data updates.

The pseudocode for the algorithm is given below.

Input: $X = \{x_1, x_2. ....x_i\}D^{n \times p}$, the number of clusters c, fuzziness parameter m
Output: Membership degree U matrix and cluster center matrix V

Begin
  1. Compute $d(V_i, V_j)$
  2. Chose $\min\{d(V_i, V_j)\}$
  3. Compute d_{ik}
  4. **If** $d_{ik} > \min\{d(V_i, V_j)\}$, then
  5. Compute $(c + 1)$ or $(c - 1) \leftarrow XB$
  6. **else**
7. Compute (c)
8. **end if**
End

4.1.2 Phase-II

The phase involves classifying the clusters obtained from phase-I into different segments based on the buying pattern of customers. One of the most effective customer segmentation methods is the RFM model (Tavakoli et al. 2018). Using the model classifies the clusters into different segments based on the above or below-average RFM attributes (Olson and Chae 2012). The list of different combinations for three-dimensional RFM attributes $(2 \times 2 \times 2)$ and their segments are given in Table 3.

4.1.3 Phase-III

In this phase, marketing strategies are devised for each segment based on the changes identified in the segments. The targeting strategies are developed based on the 4P's of the marketing mix (Kotler 2009).

5 Experiments

Since the existing dFCM algorithm had some issues, we modified them to overcome them. To assess the performance of the MdFCM algorithm, we conducted an extensive experiment using the retail supermarket dataset with eleven new data updates.

Table 3 RFM patterns

Pattern	Recency	Frequency	Monetary	Customer segments
1	High	High	High	Best
2	High	Low	High	Valuable
3	High	High	Low	Shoppers
4	High	Low	Low	First-time
5	Low	High	High	Churn
6	Low	High	Low	Frequent
7	Low	Low	High	Spenders
8	Low	Low	Low	Uncertain

5.1 Dataset description

The dataset was collected from a retail supermarket in India. The dataset contains 35,248 customer transaction records with three-dimensional RFM attributes for one year. The dataset was chosen based on its suitability for dynamic clustering algorithms, i.e., cycles starting from 01/01/2006 to 30/12/2006. The dataset can be downloaded from the R programming library. The URL for the same is given here (RFM—Customer Level Data (r-project.org)). The dataset was segregated by month, and then the RFM values were calculated, as explained in Table 4.

Below is a detailed description of the initial dataset and its subsequent new data updates.

5.1.1 Initial dataset

January 2006 is an initial dataset to cluster the new data updates. The initial period consists of 796 observations.

5.1.2 New data updates

After clustering the initial dataset, there are 11 cycles as new data updates for each month starting from February 2006 to December 2006. At the end of each month, the latest data updates are clustered to update the customer segments with new information. Then changes occurring in the customer segments are reported to managers for them to act upon. The details relating to each of these cycles are given in Table 5.

5.2 Pre-processing

Since the unit of measurement for all the three attributes is different, the initial dataset and its subsequent new data

Table 4 RFM description

Variable	Description
R	It represents the interval between the customer's last date of purchase in a month and the last date of a month. The lower recency value indicates a higher recency rate
F	The number of transactions made by customers within a month. Higher frequency implies the best frequency rate
M	The total amount of money customers has spent in a month. A higher monetary value implies the best monetary rate

Table 5 New data updates

Cycle	No. of data	Total data
February	811	1607
March	1182	2789
April	1564	4353
May	2416	6769
June	3244	10,013
July	4593	14,606
August	5020	19,626
September	5038	24,664
October	4599	29,263
November	3585	32,848
December	2400	35,248

updates are normalized between 0 to 1, keeping each month of updates separately (see Eq. 8, 9, and 10). This normalization aims to facilitate the smooth classification of the clusters into various segments as part of the phase-II DCS framework. For example, if the obtained cluster centers are above 0.5, they will be considered above-average ($\uparrow$). In the same way, if the obtained cluster centers are below 0.5, they will be regarded as below-average ($\downarrow$).

$$\text{Recency}: \quad 1 - \frac{R_k}{Max\{R_k\}} \quad \text{where} \quad k = n+1, \ldots, n+m \tag{8}$$

$$\text{Frequency}: \quad \frac{F_k}{Max\{F_k\}} \quad \text{where} \quad k = n+1, \ldots, n+m \tag{9}$$

$$\text{Monetary}: \frac{M_k}{Max\{M_k\}} \quad \text{where} \quad k = n+1, \ldots, n+m. \tag{10}$$

5.3 Experimental environment

The experiments were conducted using Intel Core i3 CPU M380 at 2.53 GHz with 4 GB of memory in the Windows 10 environment. We have used "R Studio" with the "fclust" library to perform the clustering task. We have used the "RcmdrPlugin.FuzzyClust" library to check the statistical significance of clusters.

5.3.1 Parameter settings

There are clustering parameters that need to be set by the user for both existing and proposed algorithms. To make the comparison fair, we have used the default parameters of these algorithms (Fathabadi 2016), as given in Table 6.

5.4 Performance evaluation

We are using some of the widely used fuzzy clustering evaluation measures to compare the performance of the MdFCM with dFCM algorithm. The following section explains these measures in detail.

5.4.1 Partition coefficient (PC)

The index measures the overall strength of membership values of fuzzy partition (Bezdek 1973). The index is defined in Eq. (11). A cluster with a maximum value is considered an optimal clustering result.

$$PC = \frac{1}{n+m} \sum_{i=1}^{c} \sum_{k=n+1}^{m} \mu_{ij}^2. \tag{11}$$

5.4.2 Modified partition coefficient (MPC)

It is introduced by Dave (Dave 1996) to solve the problem of monotonic evolution tendency with c of PC index. The

Table 6 Parameter settings

Algorithm	Parameters	Values
dFCM	E_{FCM}	0.01
	Y_{thr}	0.8
	z	2
MdFCM	E_{FCM}	0.01
	z	2

index is defined in Eq. (12) and a cluster with maximum value is considered to be an optimal clustering results.

$$MPC = 1 - \frac{c}{c-1}(1 - PC). \tag{12}$$

5.4.3 Partition entropy (PE)

The index is a scalar measure of the amount of fuzziness in a fuzzy partition (Bezdek 2013). The index is defined in Eq. (13). The index with a minimum value is considered to be an effective result.

$$PE = -\frac{1}{n+m}\sum_{i=1}^{c}\sum_{k=n+1}^{m}\mu_{ij}\log_a\mu_{ij}. \tag{13}$$

5.4.4 XBI

The index measures the compactness and separateness of fuzzy partition (Xie and Beni 1991). The index with a minimum value is considered to be an effective result. The index is defined in Eq. (5).

5.4.5 Fuzzy silhouette (FS) index

The index is an extension of the crisp silhouette introduced to evaluate the fuzzy clustering algorithms (Campello and Hruschka 2006). The index with maximum value is considered to be a better result. The index is defined in Eq. (14).

$$FS = \frac{\sum_{j=1}^{n}\left(\mu_{pj} - \mu_{qj}\right)^{\alpha}S_j}{\sum_{j=1}^{n}\left(\mu_{pj} - \mu_{qj}\right)^{\alpha}}. \tag{14}$$

5.4.6 Statistical significance test

To check whether the clusters obtained from the MdFCM algorithm are statistically significant or not, we conducted MANOVA Pillai's trace statistics test (Pillai 1955). The test is defined in Eq. (15).

$$PT = \text{trace}\left[H(H+E)^{-1}\right] = \sum_{i=1}^{q}\frac{\lambda_i}{1+\lambda_i} \tag{15}$$

6 Results and discussion

The experimental results of both dFCM and MdFCM algorithms are reported in this section. The objective of the section is to demonstrate the effectiveness of the proposed MdFCM algorithm compared to the existing dFCM algorithm.

6.1 Performance of MdFCM algorithm

As the first step in dynamic clustering, we have initial clusters to cluster the new data updates. In this connection, the initial clusters are determined as five clusters based on the minimum XBI value for the initial dataset of January 2006 (see Table 7).

As part of phase-II, the initial cluster solution of five clusters is classified into different segments based on their RFM patterns (see Table 8). The first cluster has the best customers and is most valuable to the company, i.e., RFM patterns are above average. The second and third clusters have new, first-time customers, i.e., frequency is below average. The fourth cluster has churn customers who are about to leave the company, i.e., recency is below average. The fifth cluster has uncertain customers who are not so valuable to the company, i.e., RFM patterns are below average. Also, the percentage of customers present in each segment is calculated from the clusters' size to understand the segments' targeting potential.

There are four segments: best, first-time, churn, and uncertain in the initial period. To update these segments with new information and to study the changes happening in these segments, the latest data updates are subsequently clustered. According to the steps of the MdFCM algorithm, the condition (Eq. 7) is checked. Then, based on its outcome, the relevant function of creation, movement, and elimination of clusters are done to calculate their respective XBI value. The results obtained in this process for all the 11 cycles are presented in Table 9.

Now the results of the existing dFCM algorithm are obtained by checking 'Y_{thr}' and the results obtained in this process are given in Table 10. It is to be noted that the

Table 7 Initial clusters

Clusters	XBI value
2	0.2432
3	0.2851
4	0.2268
5	**0.1621**
6	0.1630
7	0.1896
8	0.1840
9	0.1697
10	0.1811
11	0.1790
12	0.1898
13	0.2070
14	0.1780
15	0.2120

The best index value is boldfaced

Table 8 Initial customer segments

Clusters	Recency	Frequency	Monetary	Size	RFM Pattern	Segments	%
1	0.8066	0.6364	0.5856	155	R↑F↑M↑	Best	19.47
2	0.8207	0.3735	0.3131	166	R↑F↓M↓	First-time	**42.59**
3	0.5069	0.4672	0.4035	173	R↑F↓M↓	First-time	
4	0.2524	0.6283	0.5852	148	R↓F↑M↑	Churn	18.59
5	0.1935	0.3778	0.3137	154	R↓F↓M↓	Uncertain	19.35
Total				796			

Maximum percentage of customers is boldfaced

number of clusters remains unchanged during the entire year of the analysis. This kind of result would yield bad customer segmentation results.

To show the effectiveness of the proposed MdFCM algorithm, its performance is compared with the performance of the existing dFCM algorithm using the fuzzy clustering evaluation measures. The comparative results of both MdFCM and dFCM (Fathabadi 2016) algorithms are presented in Table 11.

From Table 11, it can be observed that the results of the MdFCM algorithm are different from the May cycle onwards. The fuzzy clustering evaluation measures indicate that the MdFCM algorithm gives effective clustering results. To test the statistical significance of the results given by the modified algorithm, the MANOVA test is conducted. The statistical significance results are given in Table 12.

From Table 12, it can be observed that the results produced by the modified algorithm are statistically significant for all the cycles. Since the MdFCM algorithm gives much better clustering results, we went ahead with it for conducting the case study on the retail supermarket.

Table 10 Results of dFCM algorithm

Cycle	$> Y_{thr}$	XBI			Total clusters
		c	$c+1$	$c-1$	
February	189	**0.1607**	0.1738	0.2011	5
March	225	**0.1719**	0.1881	0.2196	5
April	306	**0.1712**	0.1930	0.2250	5
May	426	**0.1737**	0.2047	0.2103	5
June	663	**0.1829**	0.1922	0.2055	5
July	928	**0.1905**	0.2171	0.2078	5
August	1006	**0.1915**	0.2210	0.2566	5
September	1004	**0.1949**	0.3500	0.2263	5
October	771	**0.1910**	0.2513	0.3261	5
November	754	**0.1890**	0.3692	0.2605	5
December	467	**0.1909**	0.2987	0.3265	5

The best index is boldfaced

7 Case study

This section presents a real-life case study that demonstrates the utility of the proposed DCS framework in solving the DCS problem. To achieve that objective, we

Table 9 The result of the MdFCM algorithm

Cycle	Condition satisfied	Count	XBI			Total clusters
			c	$c+1$	$c-1$	
February	Yes	14	0.1648	**0.1607**	0.2679	5
March	Yes	37	0.1835	**0.1719**	0.2765	5
April	Yes	37	0.1762	**0.1712**	0.2848	5
May	Yes	66	**0.1634**	0.1736	0.2587	6
June	Yes	113	0.2174	**0.1730**	0.1829	6
July	Yes	146	0.2101	**0.1823**	0.1905	6
August	Yes	192	**0.1767**	0.1803	0.1915	7
September	Yes	216	**0.1779**	0.1784	0.1855	8
October	Yes	255	0.1951	**0.1713**	0.1806	8
November	Yes	133	0.1971	**0.1769**	0.1819	8
December	Yes	125	0.1951	0.1793	**0.1722**	7

The best index values are boldfaced

Table 11 Clustering validation results

Cycle	PC		MPC		PE		FS index		XBI	
	dFCM	MdFCM	dFCM	MdFCM	dFCM	MdFCM	dFCM	MdFCM	dFCM	MdFCM
February	0.4856	0.4856	0.3570	0.3570	1.0320	1.0320	0.6222	0.6222	0.1607	0.1607
March	0.4850	0.4850	0.3563	0.3563	1.0305	1.0305	0.6163	0.6163	0.1719	0.1719
April	0.4861	0.4861	0.3576	0.3576	1.0278	1.0278	0.6175	0.6175	0.1712	0.1712
May	0.4411	**0.4820**	0.3293	**0.3526**	1.1691	**1.0347**	0.5984	**0.6129**	0.1737	**0.1634**
June	0.4423	**0.4823**	0.3307	**0.3528**	1.1649	**1.0330**	0.5962	**0.6129**	0.1829	**0.1730**
July	0.4449	**0.4837**	0.3339	**0.3546**	1.1583	**1.0293**	0.5966	**0.6144**	0.1905	**0.1823**
August	0.4149	**0.4856**	0.3174	**0.3570**	1.2648	**1.0247**	0.5792	**0.6147**	0.1915	**0.1767**
September	0.3899	**0.4856**	0.3027	**0.3571**	1.3597	**1.0240**	0.5636	**0.6138**	0.1949	**0.1779**
October	0.3890	**0.4847**	0.2900	**0.3558**	1.3621	**1.0261**	0.5619	**0.6124**	0.1910	**0.1713**
November	0.3888	**0.4857**	0.3015	**0.3571**	1.3623	**1.0245**	0.5612	**0.6145**	0.1890	**0.1769**
December	0.3886	**0.4854**	0.3172	**0.3568**	1.2655	**1.0249**	0.5805	**0.6138**	0.1909	**0.1722**

Best results are boldfaced

Table 12 Result of statistical test

Cycle	df	Pillai	Approx F	Numerator Df	Denominator Df	Pr (> F)	Residuals
February	4	1.4864	393.2883	12	4806	**0.0000**	1602
March	4	1.4783	676.1501	12	8352	**0.0000**	2784
April	4	1.4785	1056.29	12	13,044	**0.0000**	4348
May	5	1.5336	1414.677	15	20,289	**0.0000**	6763
June	5	1.5237	2065.592	15	30,021	**0.0000**	10,007
July	5	1.5233	3012.054	15	43,800	**0.0000**	14,600
August	6	1.5537	3512.526	18	58,857	**0.0000**	19,619
September	7	1.5983	4016.235	21	73,968	**0.0000**	24,656
October	7	1.6035	4798.676	21	87,765	**0.0000**	29,255
November	7	1.6106	5438.302	21	98,520	**0.0000**	32,840
December	6	1.5555	6325.133	18	105,723	**0.0000**	35,241

Significant results at 0.05 levels are boldfaced

have developed a case concerning a retail supermarket located in southern India. For confidentiality reasons, they requested to be anonymous. The supermarket has a wide variety of products on offer to its customers. The decision-makers of this supermarket wanted to segment their large customer base based on the buying pattern of customers so that they could target their customers with appropriate marketing strategies. In particular, they wanted to adjust their marketing strategies from time to time based on the changes identified in their segments by clustering the new data updates continuously. In doing so, they are trying to optimize their resources to gain more profit and customer satisfaction. The following section elaborates on the case study conducted.

7.1 The DCS results

The data collected during the months starting from February 2006 to December 2006 was used to update the customer segments (see Table 8) with new information according to the process described in Fig. 1. The updated customer segments are presented in Table 13; as can be seen in Table 13 that the maximum percentage of customers are present in the first-time segment throughout the year.

The marketing strategies devised for each customer segment are presented in Table 14. Here, the percentage of customers present in each segment was used as a reference point to adjust the marketing strategies from time to time.

Table 13 Results of DCS

Cycle	Clusters	Recency	Frequency	Monetary	Size	RFM Pattern	Segments	Customer (%)
February	1	0.7747	0.6039	0.5830	298	R↑F↑M↑	Best	18.54
	2	0.5381	0.4068	0.3742	383	R↑F↓M↓	First-time	**44.31**
	3	0.8677	0.3612	0.3303	329	R↑F↓M↓	First-time	
	4	0.2424	0.5773	0.5588	292	R↓F↑M↑	Churn	18.17
	5	0.1974	0.3548	0.3168	305	R↓F↓M↓	Uncertain	18.98
	Total				1607			
March	1	0.7717	0.5905	0.5688	528	R↑F↑M↑	Best	18.93
	2	0.8689	0.3744	0.3392	565	R↑F↓M↓	First-time	**43.89**
	3	0.5420	0.3956	0.3577	659	R↑F↓M↓	First-time	
	4	0.2358	0.5749	0.5549	496	R↓F↑M↑	Churn	17.78
	5	0.1938	0.3631	0.3245	541	R↓F↓M↓	Uncertain	19.40
	Total				2789			
April	1	0.7796	0.5618	0.5532	827	R↑F↑M↑	Best	19.00
	2	0.8660	0.3591	0.3329	919	R↑F↓M↓	First-time	**44.50**
	3	0.5370	0.3742	0.3440	1018	R↑F↓M↓	First-time	
	4	0.2514	0.5554	0.5438	761	R↓F↑M↑	Churn	17.48
	5	0.1930	0.3469	0.3162	828	R↓F↓M↓	Uncertain	19.02
	Total				4353			
May	1	0.8450	0.5596	0.5341	1079	R↑F↑M↑	Best	31.08
	2	0.5224	0.5488	0.5218	1025	R↑F↑M↑	Best	
	3	0.8678	0.3627	0.3251	1263	R↑F↓M↓	First-time	**37.15**
	4	0.5441	0.3530	0.3103	1252	R↑F↓M↓	First-time	
	5	0.1875	0.5501	0.5234	956	R↓F↑M↑	Churn	14.12
	6	0.1951	0.3457	0.3074	1194	R↓F↓M↓	Uncertain	17.64
	Total				6769			
June	1	0.8606	0.5317	0.5130	1519	R↑F↑M↑	Best	30.92
	2	0.5591	0.5220	0.5016	1577	R↑F↑M↑	Best	
	3	0.8547	0.3443	0.3092	1886	R↑F↓M↓	First-time	**37.61**
	4	0.5075	0.3424	0.3048	1880	R↑F↓M↓	First-time	
	5	0.2039	0.5409	0.5194	1444	R↓F↑M↑	Churn	14.42
	6	0.1796	0.3459	0.3090	1707	R↓F↓M↓	Uncertain	17.05
	Total				10,013			
July	1	0.8374	0.5396	0.5130	2341	R↑F↑M↑	Best	16.03
	2	0.5014	0.5217	0.4925	2279	R↑F↑M↓	Shoppers	15.60
	3	0.8660	0.3567	0.3137	2762	R↑F↓M↓	First-time	**37.05**
	4	0.5434	0.3396	0.2953	2650	R↑F↓M↓	First-time	
	5	0.1797	0.5276	0.4971	2037	R↓F↑M↓	Frequent	13.95
	6	0.1800	0.3407	0.2958	2537	R↓F↓M↓	Uncertain	17.37
	Total				14,606			
August	1	0.8471	0.5445	0.5198	2710	R↑F↑M↑	Best	27.44
	2	0.5172	0.5416	0.5162	2675	R↑F↑M↑	Best	
	3	0.8963	0.3626	0.3224	2932	R↑F↓M↓	First-time	**30.57**
	4	0.6637	0.3587	0.3154	3068	R↑F↓M↓	First-time	
	5	0.1794	0.5345	0.5073	2512	R↓F↑M↑	Churn	12.80
	6	0.4059	0.3577	0.3149	2985	R↓F↓M↓	Uncertain	29.19
	7	0.1447	0.3509	0.3067	2744	R↓F↓M↓	Uncertain	
	Total				19,626			

Table 13 (continued)

Cycle	Clusters	Recency	Frequency	Monetary	Size	RFM Pattern	Segments	Customer (%)
September	1	0.7259	0.5842	0.5482	2634	R↑F↑M↑	Best	10.68
	2	0.8937	0.4786	0.4403	3101	R↑F↓M↓	First-time	**40.16**
	3	0.8667	0.3290	0.2811	3081	R↑F↓M↓	First-time	
	4	0.6195	0.3671	0.3192	3724	R↑F↓M↓	First-time	
	5	0.1596	0.5359	0.4989	2940	R↓F↑M↓	Frequent	23.48
	6	0.4383	0.5276	0.4907	2850	R↓F↑M↓	Frequent	
	7	0.3688	0.3533	0.3043	3305	R↓F↓M↓	Uncertain	25.68
	8	0.1317	0.3529	0.3030	3029	R↓F↓M↓	Uncertain	
	Total				24,664			
October	1	0.6960	0.5863	0.5567	3076	R↑F↑M↑	Best	10.51
	2	0.8950	0.4952	0.4607	3582	R↑F↓M↓	First-time	**39.68**
	3	0.8715	0.3350	0.2878	3678	R↑F↓M↓	First-time	
	4	0.6351	0.3795	0.3334	4353	R↑F↓M↓	First-time	
	5	0.1522	0.5381	0.5065	3458	R↓F↑M↑	Churn	11.82
	6	0.4104	0.5228	0.4896	3520	R↓F↑M↓	Frequent	12.03
	7	0.3884	0.3450	0.2984	3705	R↓F↓M↓	Uncertain	25.96
	8	0.1373	0.3550	0.3071	3891	R↓F↓M↓	Uncertain	
	Total				29,263			
November	1	0.6521	0.5467	0.5139	3885	R↑F↑M↑	Best	11.83
	2	0.8865	0.5276	0.4955	3967	R↑F↑M↓	Shoppers	12.08
	3	0.8860	0.3541	0.3067	4549	R↑F↓M↓	First-time	**26.85**
	4	0.6404	0.3573	0.3090	4270	R↑F↓M↓	First-time	
	5	0.3583	0.5774	0.5478	3357	R↓F↑M↑	Churn	10.22
	6	0.1393	0.5056	0.4686	3975	R↓F↑M↓	Frequent	12.10
	7	0.4037	0.3792	0.3328	4821	R↓F↓M↓	Uncertain	26.93
	8	0.1526	0.3365	0.2869	4024	R↓F↓M↓	Uncertain	
	Total				32,848			
December	1	0.8564	0.5436	0.5102	4958	R↑F↑M↑	Best	27.09
	2	0.5451	0.5618	0.5306	4590	R↑F↑M↑	Best	
	3	0.8817	0.3658	0.3175	5181	R↑F↓M↓	First-time	**29.99**
	4	0.6229	0.3603	0.3112	5391	R↑F↓M↓	First-time	
	5	0.1738	0.5523	0.5179	4862	R↓F↑M↑	Churn	13.79
	6	0.3751	0.3981	0.3501	5292	R↓F↓M↓	Uncertain	29.13
	7	0.1429	0.3552	0.3054	4974	R↓F↓M↓	Uncertain	
	Total				35,248			

Maximum percentage is boldfaced

7.2 Managerial implications

The study results give important managerial insights. The retail industry is flooded with many companies; hence, heavy competition exists. Because of the competition, each player is deploying innovative marketing strategies to win their customers. Central to this task is machine learning algorithms that can help understand the customers to the minute level. In this direction, we segmented the customers into best, shoppers, first-time, churn, frequent, and uncertain. Throughout the year, the first-time segment had the maximum number of customers; this implies that the company could not convert first-time customers into more profitable ones. Pursuing this segment of customers would improve the wealth of the company. The study results can be used as a useful reference for the retail industry.

Table 14 Marketing strategies

Cycle	Product	Price	Place	Promotion	Remarks
February		Offer discount			Focus on first-time customers
March		Offer discount		Social media outreach	Focus on reaching out first-time customers
April	Offer free products				Focus on product bundling
May			Offer the products online		Encourage online buying
June		Customer retaining strategies		Distribute coupons	Primarily focus on best customers
July	Innovative packaging			Start the advertising campaigns	Focus on shoppers and frequent customers
August	Discount on seasonal products			Give free samples	Focus on uncertain customers
September		Offer percentage based discount	Instant delivery		Retain frequent customers
October	Offer product trials	Discount	Free delivery	Email marketing	Pursue the first-time customers
November		Implement competitive pricing strategies			Focus on convincing the customers
December	Implement loyalty programs		Offer product bundles		Focus on retaining the customers

8 Conclusion

Applying machine learning techniques to solve real-world problems is important in the growing trend of quantitative decision-making. In line with this, the study has solved a real-world problem faced by the retail supermarket. In comparison with the previous literature, the major contributions of our study are summarized as follows: (1) We proposed a DCS framework that can be used as a tool by managers in doing DCS (2) To overcome the shortcoming of the existing dFCM algorithm, it has been modified, resulting in the MdFCM algorithm. To prove the effectiveness of the modified algorithm, extensive experiments were conducted using the retail supermarket dataset with eleven cycles. Experimental results were then compared with the existing algorithm to show that the proposed algorithm performs better than the existing one. Once the effectiveness of the proposed algorithm was proved, it was employed to conduct the case study on a retail supermarket. The case study discussed the various customer segments of the company and suggested successful marketing strategies to deal with them. In future work, interval-valued fuzzy sets can be used for DCS.

Declarations

References

Allaway AW, D'Souza G, Berkowitz D, Kim KK (2014) Dynamic segmentation of loyalty program behavior. J Mark Anal 2:18–32. https://doi.org/10.1057/jma.2014.2

Anitha P, Patil MM (2020) RFM model for customer purchase behavior using K-means algorithm. J King Saud Univ Comput Inf Sci. https://doi.org/10.1016/j.jksuci.2019.12.011

Arunachalam D, Kumar N (2018) Benefit-based consumer segmentation and performance evaluation of clustering approaches: an evidence of data-driven decision-making. Expert Syst Appl 111:11–34. https://doi.org/10.1016/j.eswa.2018.03.007

Askari S (2021) Noise-resistant fuzzy clustering algorithm. Granul Comput 6:815–828. https://doi.org/10.1007/s41066-020-00230-6

Bachtiar FA (2019) Customer segmentation using two-step mining method based on RFM model. Int Conf Sustain Inf Eng Technol 2018:10–15. https://doi.org/10.1109/siet.2018.8693173

Bansal HS, Taylor SF, James YS (2005) "Migrating" to new service providers: toward a unifying framework of consumers' switching behaviors. J Acad Mark Sci 33:96–115. https://doi.org/10.1177/0092070304267928

Bernstein F, Modaresi S, Sauré D (2016) A dynamic clustering approach to data-driven assortment personalization. SSRN J. https://doi.org/10.2139/ssrn.2983207

Bezdek JC (1973) Cluster validity with fuzzy sets. J Cybern 3:58–73. https://doi.org/10.1080/01969727308546047

Bezdek JC (2013) Pattern recognition with fuzzy objective function algorithms. Springer Science & Business Media, Berlin

Bezdek JC, Ehrlich R, Full W (1984) FCM: the fuzzy *c*-means clustering algorithm. Comput Geosci 10:191–203. https://doi.org/10.1016/0098-3004(84)90020-7

Bose I, Chen X (2014) Detecting temporal changes in customer behavior. Int Electr Eng Congr iEECON 2014:3–6. https://doi.org/10.1109/iEECON.2014.6925923

Bose I, Chen X (2015) Detecting the migration of mobile service customers using fuzzy clustering. Inf Manag 52:227–238. https://doi.org/10.1016/j.im.2014.11.001

Cachon GP (2012) Dynamic versus static pricing in the presence of strategic consumers. P 1–31

Campello RJGB, Hruschka ER (2006) A fuzzy extension of the silhouette width criterion for cluster analysis. Fuzzy Sets Syst 157:2858–2875. https://doi.org/10.1016/j.fss.2006.07.006

Chan CCH, Bin CC, Hsien WC (2011) Pricing and promotion strategies of an online shop based on customer segmentation and multiple objective decision making. Expert Syst Appl 38:14585–14591. https://doi.org/10.1016/j.eswa.2011.05.024

Chen SM (1997) Interval-valued fuzzy hypergraph and fuzzy partition. IEEE Trans Syst Man Cybern Part B Cybern 27:725–733. https://doi.org/10.1109/3477.604121

Chen SM, Hsiao WH (2000) Bidirectional approximate reasoning for rule-based systems using interval-valued fuzzy sets. Fuzzy Sets Syst 113:185–203. https://doi.org/10.1016/S0165-0114(98)00351-0

Chen SM, Hsiao WH, Jong WT (1997) Bidirectional approximate reasoning based on interval-valued fuzzy sets. Fuzzy Sets Syst 91:339–353. https://doi.org/10.1016/S0165-0114(97)86594-3

Dave RN (1996) Validating fuzzy partitions obtained through c-shells clustering. Pattern Recognit Lett 17:613–623. https://doi.org/10.1016/0167-8655(96)00026-8

Fan B, Zhang P (2009) Spatially enabled customer segmentation using a data classification method with uncertain predicates. Decis Support Syst 47:343–353. https://doi.org/10.1016/j.dss.2009.03.002

Fathabadi H (2016) Power distribution network reconfiguration for power loss minimization using novel dynamic fuzzy *c*-means (dFCM) clustering based ANN approach. Int J Electr Power Energy Syst 78:96–107. https://doi.org/10.1016/j.ijepes.2015.11.077

Garai T, Chakraborty D, Roy TK (2019) A fuzzy rough multi-objective multi-item inventory model with both stock-dependent demand and holding cost rate. Granul Comput 4:71–88. https://doi.org/10.1007/s41066-018-0085-6

Gür Ali Ö, Aritürk U (2014) Dynamic churn prediction framework with more effective use of rare event data: the case of private banking. Expert Syst Appl 41:7889–7903. https://doi.org/10.1016/j.eswa.2014.06.018

Ha SH, Bae SM, Park SC (2002) Customer's time-variant purchase behavior and corresponding marketing strategies: an online retailer's case. Comput Ind Eng 43:801–820. https://doi.org/10.1016/S0360-8352(02)00141-9

Han SH, Lu SX, Leung SCH (2012) Segmentation of telecom customers based on customer value by decision tree model. Expert Syst Appl 39:3964–3973. https://doi.org/10.1016/j.eswa.2011.09.034

Hiziroglu A (2013) Soft computing applications in customer segmentation: state-of-art review and critique. Expert Syst Appl 40:6491–6507. https://doi.org/10.1016/j.eswa.2013.05.052

Hu X, Zhao S (2015) A dynamic segmentation method of power customer based on rough clustering. Chinese Control Conf CCC 2015-Septe:8773–8778. https://doi.org/10.1109/ChiCC.2015.7261026

Hughes A (1994) Strategic database marketing. Probus Publ, Chicago

Khan R, Lewis M, Singh V (2009) Dynamic customer management and the value of one-to-one marketing. Mark Sci 28:1063–1079. https://doi.org/10.1287/mksc.1090.0497

Kim SY, Jung TS, Suh EH, Hwang HS (2006) Customer segmentation and strategy development based on customer lifetime value: a case study. Expert Syst Appl 31:101–107. https://doi.org/10.1016/j.eswa.2005.09.004

Kotler P (2009) Marketing management: a south Asian perspective. Pearson Education India, Noida

Lim S, Lee B (2015) Loyalty programs and dynamic consumer preference in online markets. Decis Support Syst 78:104–112. https://doi.org/10.1016/j.dss.2015.05.008

Lingras P, Hogo M, Snorek M, West C (2005) Temporal analysis of clusters of supermarket customers: Conventional versus interval set approach. Inf Sci (NY) 172:215–240. https://doi.org/10.1016/j.ins.2004.12.007

Macqueen J (1967) Some methods for classification and analysis of multivariate observations. In: Proceedings of the fifth Berkeley symposium on mathematical statistics and probability, vol 1, p 281–297. citeulike-article-id:6083430

Nadi S, Saraee M, Davarpanah-Jazi M (2010) A fuzzy recommender system for dynamic prediction of user's behavior. In: 2010 International conference for internet technology and secured transactions (ICITST)

Olson DL, Chae B (2012) Direct marketing decision support through predictive customer response modeling. Decis Support Syst 54:443–451. https://doi.org/10.1016/j.dss.2012.06.005

Peters G, Weber R (2012) Dynamic clustering with soft computing. Wiley Interdiscip Rev Data Min Knowl Discov 2:226–236. https://doi.org/10.1002/widm.1050

Peters G, Weber R, Nowatzke R (2012) Dynamic rough clustering and its applications. Appl Soft Comput J 12:3193–3207. https://doi.org/10.1016/j.asoc.2012.05.015

Pillai KCS (1955) Some new test criteria in multivariate analysis. Ann Math Stat. https://doi.org/10.1214/aoms/1177728599

Reutterer T, Mild A, Natter M, Taudes A (2006) A dynamic segmentation approach for targeting and customizing direct marketing campaigns. J Interact Mark. https://doi.org/10.1002/dir.20066

Rust RT, Kumar V, Venkatesan R (2011) Will the frog change into a prince? Predicting future customer profitability. Int J Res Mark 28:281–294. https://doi.org/10.1016/j.ijresmar.2011.05.003

Seret A, Vanden Broucke SKLM, Baesens B, Vanthienen J (2014) A dynamic understanding of customer behavior processes based on clustering and sequence mining. Expert Syst Appl 41:4648–4657. https://doi.org/10.1016/j.eswa.2014.01.022

Smith WR (1956) Product differentiation and market segmentation as alternative marketing strategies. J Mark 21:3. https://doi.org/10.2307/1247695

Tavakoli M, Molavi M, Masoumi V et al (2018) Customer segmentation and strategy development based on user behavior analysis, RFM model and data mining techniques: a case study. In: Proceedings—15th international conference on e-business engineering, ICEBE, p 119–126. https://doi.org/10.1109/ICEBE.2018.00027

Teichert T, Shehu E, von Wartburg I (2008) Customer segmentation revisited: the case of the airline industry. Transp Res Part A Policy Pract 42:227–242. https://doi.org/10.1016/j.tra.2007.08.003

Turksen IB (1986) Interval valued fuzzy sets based on normal forms. Fuzzy Sets Syst 20:191–210. https://doi.org/10.1016/0165-0114(86)90077-1

Wedel M, Kamakura WA (2000) Market segmentation: conceptual and methodological foundations (international series in quantitative marketing)

Wilson J, Chaudhury S, Lall B (2018) Clustering short temporal behaviour sequences for customer segmentation using LDA. Expert Syst 35:1–16. https://doi.org/10.1111/exsy.12250

Xie XL, Beni G (1991) A validity measure for fuzzy clustering. IEEE Trans Pattern Anal Mach Intell 13:841–847

Yao Z (2013) Visual customer segmentation and behavior analysis—a SOM-based approach

You Z, Si YW, Zhang D et al (2015) A decision-making framework for precision marketing. Expert Syst Appl 42:3357–3367. https://doi.org/10.1016/j.eswa.2014.12.022

Zeybek H (2018) Customer segmentation strategy for rail freight market: the case of Turkish State Railways. Res Transp Bus Manag. https://doi.org/10.1016/j.rtbm.2018.10.003

Zhang J, Lei L, Zhang S, Song L (2017) Dynamic vs. static pricing in a supply chain with advertising. Comput Ind Eng 109:266–279. https://doi.org/10.1016/j.cie.2017.05.006

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.