

# INTEGRATED APPROACH OF FUZZY MULTI-ATTRIBUTE DECISION MAKING AND DATA MINING FOR CUSTOMER SEGMENTATION

**Manidatta Ray<sup>1</sup>, Mamata Ray<sup>2</sup>, Kamalakanta Muduli<sup>3</sup>,  
Audrius Banaitis<sup>4</sup>, Anil Kumar<sup>5</sup>**

<sup>1</sup> Birla Global University, Birla School of Management, India, ORCID: 0000-0002-5614-6469, manidatta.ray@bgu.ac.in;

<sup>2</sup> Biju Patnaik University of Technology, India, ORCID: 0000-0002-5057-8009, mamata.ray@gmail.com;

<sup>3</sup> The Papua New Guinea University of Technology, Papua New Guinea, ORCID: 0000-0002-4245-9149, kamalakantam@gmail.com;

<sup>4</sup> Vilnius Gediminas Technical University, Faculty of Civil Engineering, Department of Construction Management and Real Estate, Lithuania, ORCID: 0000-0002-3302-1209, audrius.banaitis@vilniustech.lt;

<sup>5</sup> London Metropolitan University, Guildhall School of Business and Law, United Kingdom, ORCID: 0000-0002-1691-0098, anilror@gmail.com.

**Abstract:** This research work focuses on integrating the multi attribute decision making with data mining in a fuzzy decision environment for customer relationship management. The main objective is to analyse the relation between multi attribute decision making and data mining considering a complex problem of ordering customers segments, which is based on four criteria of customer's life time value, viz. length (L), recency (R), frequency (F) and monetary value (M). The proposed integrated approach involves fuzzy C-means (FCM) cluster analysis as data mining tool. The experiment conducted using MATLAB 12.0 for identifying eight clusters of customers. The two multi attribute decision making tools i.e., fuzzy AHP (Analytic Hierarchy Process) and fuzzy TOPSIS (Technique for Order Preference by Similarity to Ideal Solution) are used for ranking these identified clusters. The applicability of the integrated decision making technique is also demonstrated in this paper considering the case of Indian retail sector. This research collected responses from nine experts from Indian retail industry regarding their perception of relative importance of four criteria of customer life value and evaluated weights of each criterion using fuzzy AHP. Transaction data of 18 months of the case retail store was analysed to segment 1,600 customers into eight clusters using fuzzy c-means clustering analysis technique. Finally, these eight clusters were ranked using fuzzy TOPSIS (Technique for Order Preference by Similarity to Ideal Solution). The findings of this research could be helpful for firms in identifying the more valuable customers for them and allocate more resources to satisfy them. The findings will be also helpful in developing different loyalty program strategies for customers of different clusters.

**Keywords:** Data mining, fuzzy c-means clustering, fuzzy AHP, customer segmentation, fuzzy TOPSIS, customer lifetime value (CLV), marketing strategies.

**JEL Classification:** C65, L81, D12.

**APA Style Citation:** Ray, M., Ray, M., Muduli, K., Banaitis, A., & Kumar, A. (2021). Integrated Approach of Fuzzy Multi-attribute Decision Making and Data Mining for Customer Segmentation. *E&M Economics and Management*, 24(4), 174–188. <https://doi.org/10.15240/tul/001/2021-4-011>

## Introduction

Recent advancements in Information and Communication Technologies (ICT) offered

various options to the consumers to select best goods and services according to their requirement. As a result, manufactures and

companies are experiencing challenges for retaining their customers and for which they are adopting innovative customer relationship management strategies. Further, it is well known by marketers that, customers have various kinds of needs and wants. Hence, they try to establish certain criteria to better identify and understand customer groups and provide preferable products and services to them in order to satisfy their varying needs and wants. In this regard segmentation is viewed by them as an important technique to create profitable segments and allocate resources to respond to the needs of selected segments based on their valuableness. However, many marketers have difficulty in identifying the right customer segments to organize marketing campaigns. This causes unsuccessful loyalty programs and promotions conjunction with waste of marketing resources. Nowadays, companies are widely employing segmentation of customers as a critical tool for differentiating them based on their buying preference (Safari et al., 2016) and understanding their behavioural responses in this competitive retail marketing era. This eventually helps the organizations in customizing their services or products as an effort to retain their customers (Safari et al., 2016; Chiang & Yang, 2018; Egemen et al., 2021; Moktadir et al., 2021). Companies can also gather the huge amount of data related to customers buying behaviours through ICT. However, it has been always a challenge for the companies to handle this massive consumer data, which is sometimes referred as Big data. To handle this massive data, Data Mining (DM) has been emerged as an effective process that extracts important information from the same through various computing, statistical and mathematical, techniques (Chiang & Yang, 2018; Guo et al., 2020). Some of the most important functions of the DM includes association, classification, clustering, prediction and visualization (Tsai, 2012). Recently, there has been increased interest in integrating operations research techniques with DM (Malik et al., 2018; Pérez-Martín et al., 2018; Liao et al., 2019; Guerard et al., 2021).

Multi-attribute decision-making (MADM) is a class of problem-solving techniques in multi-criteria decision-making (MCDM) which belongs to operations research and it deals with multiple criteria in various decision environments, viz. deterministic, stochastic and fuzzy. The literature

on MADM-DM integration shows some new methodologies in various application areas. For example, Pineda et al. (2018) proposed an integrated model that combines data mining and multi criteria decision making techniques, to identify and diagnose financial and operational performance of airlines. Amiri et al. (2021) presented a new model with a triangular fuzzy approach for sustainable supplier selection (SSS) in the supply chain. Arabameri et al. (2019) applied four bivariate and multivariate statistical (WOE and BLR), data mining (RF) and multi criteria decision making (TOPSIS) models for ground water potential mapping. Marques et al. (2020) by combining fuzzy cognitive mapping techniques and the system dynamics approach created an analysis model that allows for a more holistic perspective on determinants of family business growth, their cause-and-effect relationships, and thus their long-term behavior. Mahdiraji et al. (2019) proposed a model based on big data analysis for evaluation of marketing strategy using clustering-multi criteria decision making approach. Mohandes et al. (2020) developed a novel Risk Assessment Model (RAM) through the integration of the Fuzzy Best Worst Method (FBWM) with the Interval-Valued Fuzzy Technique for Order of Preference by Similarity to Ideal Solution (IVFTOPSIS). Liou et al. (2021) proposed a novel hybrid MCDM model, that integrates the support vector machine (SVM), the fuzzy best worst method (FBWM) and, the fuzzy technique for order preference by similarity to an ideal solution (FTOPSIS) approaches to select the most suitable green suppliers. Ozkaya et al. (2021) used a hybrid model of data mining and multi criteria decision making methods to propose an indicator that measures science, technology and innovation (STI) policies of forty countries to reach their sustainable development goals. Ray and Mangaraj (2016) evaluated CLV in terms of length, recency, frequency, monetary (LRFM) variables and employed AHP to obtain relative weights of these variable based on multiple experts. They segmented customers into clusters by K-means algorithm using Davis-Bouldin (DB) index for a retailer to design promotional strategies for improving retailer-customers relationship. Wu and Olson (2006) developed a TOPSIS based DM technique to classify credit score data into groups of high expected repayment and low expected repayment. Aghdaie et al. (2014) proposed an approach that included K-means cluster

analysis as a DM tool along with SWARA and VIKOR as two MADM tools to rank the clusters of suppliers.

The current paper proposes a new hybrid MADM-DM approach to evaluate CLV in terms of LRFM variables in a fuzzy decision environment. For the same, AHP as well as fuzzy AHP (FAHP) are used to obtain relative weights of these variables from a group of experts. To employ fuzzy C-means (FCM) clustering to segment customers and later, rank these clusters of customers using a fuzzy TOPSIS (FTOPSIS) method. Section 1 presents a discussion about customer lifetime value or loyalty as the composite construct to segment customers and to design marketing strategies accordingly. In section 2, proposed integrated MADM-DM methodology in a fuzzy decision environment is discussed. Section 3 demonstrates the applicability of the suggested integrated DM approach in Indian retail sector. Section 4 presents a discussions and implications, and finally, the last section concludes the paper.

### 1. Customer Life-Time Value (CLV)

Customer value analysis is a type of analytic tool to understand a customer from behavioural point of view. According to Kotler (2003), customers are broadly classified by two dimensions, viz. customer characteristics and behaviour. Demographics, psychographic and geographic variables are included in customer characteristics, whereas attitudes towards the product and the response shown by the customers explains the behavioural dimension. Customer value also termed as 'life time value (LTV)' or 'customer lifetime value (CLV)' or 'customer equity (CE)' or even 'customer profitability (CP)' is assessed by evaluating the present value of the future profit stream which is expected over a given time horizon of transacting with the customer. Wiesel et al. (2011) states it as the future cash flow value to be produced from a customer in CRM that also determines the present value of customers brought to an organization during customers' life cycle. Hence, it is normally used to recognize beneficial customers and formulate appropriate strategies for different customers' segments (Khajvand et al., 2011). However, other areas of application of CLV includes customers evaluation where the segmentation scheme helps the marketer to make a decision whether

it is better to acquire few large customers (who may be risky) or a large number of small customers (Benoit & Van den Poel, 2009); product recommendation to each customer group (Liu & Shih, 2005) etc. Literature shows that there are several modelling approaches to study CLV. These approaches include RFM models, econometric models, probability models, persistence models, computer models, diffusion/growth models etc. Measuring RFM is an important method for assessing CLV.

RFM scoring model using RFM data was proposed by Hughes (1994) which differentiated customers from a large database by these three variables, so that different marketing strategies could be adopted for different groups of customers. RFM model was also used for assessing CLV value by Safari et al. (2016). The authors further employed fuzzy clustering and fuzzy AHP approach for segmenting the customers and ranking the clusters respectively. Anitha and Patil (2019) in their research aimed at segmenting customers based on their CLV value also considered RFM variables. However, few researchers argue that RFM variable based assessment of CLV fails to identify customers based on their transactional length (Zare & Emadi, 2020). Hence, an extended RFM model was proposed by adding another dimension the customer transactional length (L) to it which is known as LRFM model (Chang & Tsay, 2004). The additional variable in LRFM model represents the time interval between the first and last exchanges with the customer. Wei et al. (2012) applied LRFM model and used self-organizing maps to segment patients and targeted important patients in a children's dental clinic. It is to be noted that, these variables remain independent in both the RFM (Liu & Shih, 2005) and LRFM (Seyed Hosseini et al., 2010) models.

Regardless of the number of variables are concerned in the CLV construct, there are two schools of thought concerning assignment of weights to variables in the modelling process. Hughes (1994) considered equal importance to R, F, M variables, and hence equal weights to them. On the other hand, some researchers have used the weighted models (Liu & Shih, 2005; Seyed Hosseini et al., 2010) in their studies. However, assignment of weights to these criteria remains a difficult task, although some authors (Liu & Shih, 2005; Rad et al., 2011; Ray & Mangaraj, 2016) have used AHP

as a systematic method for weighing these variables. Conceptually, AHP determines relative weights of alternatives with respect to one consistent expert whose subjective opinions are considered in a pair-wise comparison matrix. In the event of multiple experts, we use arithmetic mean for averaging these weights obtained from multiple comparison matrices associated with consistent experts. But in this paper, the distribution of judgements contained in those comparison matrices are accounted and represent them as a triangular fuzzy number in an aggregated comparison matrix. Hence, instead of multiple pairwise comparison matrices consisting of judgements either in a scale of 1 to 9 or an inverse scale 1 to 1/9, we get a matrix of triangular fuzzy numbers (TFNs), whose evaluation needs a FAHP for the determination of weights. The detailed procedure in the next section.

## 2. The Proposed Model

This section describes the proposed MADM-DM modelling approach for ranking customer segments using L, R, F and M scores. This requires preparation of customer CLV file for generating L, R, F and M data of each customer by accessing store's data-base for a particular time period looking for data such as card number, customer name, and date of visit to the store and purchase amount. The customer record in this respect is generated consisting of the fields, viz. card number and name of the customer along with his value of R, L, F and M with reference to a particular date. This becomes the data preparation stage which then leads to the application of the integrated model to assess customer's value for the store based on his buying behavior by employing FCM, FAHP and FTOPSIS in a sequential manner as shown in Fig. 1. Details of these models are explained in the following sub-sections.

### 2.1 Fuzzy C-Mean (FCM) Clustering

For customers grouping into segments and as a DM technique, clustering are used. With the help of clustering the natural groups can be identified from a large set of data with LTV for example K-means, kohonen network/self-organizing map and fuzzy c-means. K-means clustering is a method frequently used to categorize data into K number of groups (Rad et al., 2011). Its algorithm performs based on crisp partitioning which means each datum belongs

to just one cluster. Hence, the membership degree of each datum in a cluster is either 0 or 1. However, for a datum having ambiguous characteristics, the membership function takes a value in the range 0 to 1. Hence, the concept of fuzzy set is used in clustering for soft partitioning of the objects into groups thereby resulting in fuzzy c-means (FCM) clustering. FCM clustering was proposed by Dunn (1973) and further developed by Bezdek et al. (1984) that allowed one piece of data to belong to two or more clusters. By this method, we allow a piece of data that can exist in multiple clusters where it explains its belongingness to each cluster to a certain degree in the range 0 to 1. Hence, we minimize the following objective function:

$$J_m = \sum_{i=1}^N \sum_{j=1}^C U_{ij}^m \|x_i - C_j\|^2, \quad (1)$$

$$1 < m < \infty.$$

Here  $N$  represents the number of data set and  $C$  represents number of clusters;  $x_i$  denotes  $i^{th}$  data value,  $C_j$  symbolizes  $j^{th}$  cluster's center;  $U_{ij}$  denotes membership degree of  $x_i$  belonging to the cluster  $j$ ;  $\|*\|$  Euclidean vector norm expressing the distance between  $j^{th}$  cluster's center and  $i^{th}$  data.

But, in case of FCM clustering, an iterative optimization of the  $J_m$  is done throughout with the update of  $U_{ij}^m$  and  $C_j$  by following formulas (2) and (3) respectively.

$$U_{ij}^m = \frac{1}{\sum_{k=1}^C \left( \frac{\|x_i - C_j\|}{\|x_i - C_k\|} \right)^{\frac{2}{m-1}}}, \quad (2)$$

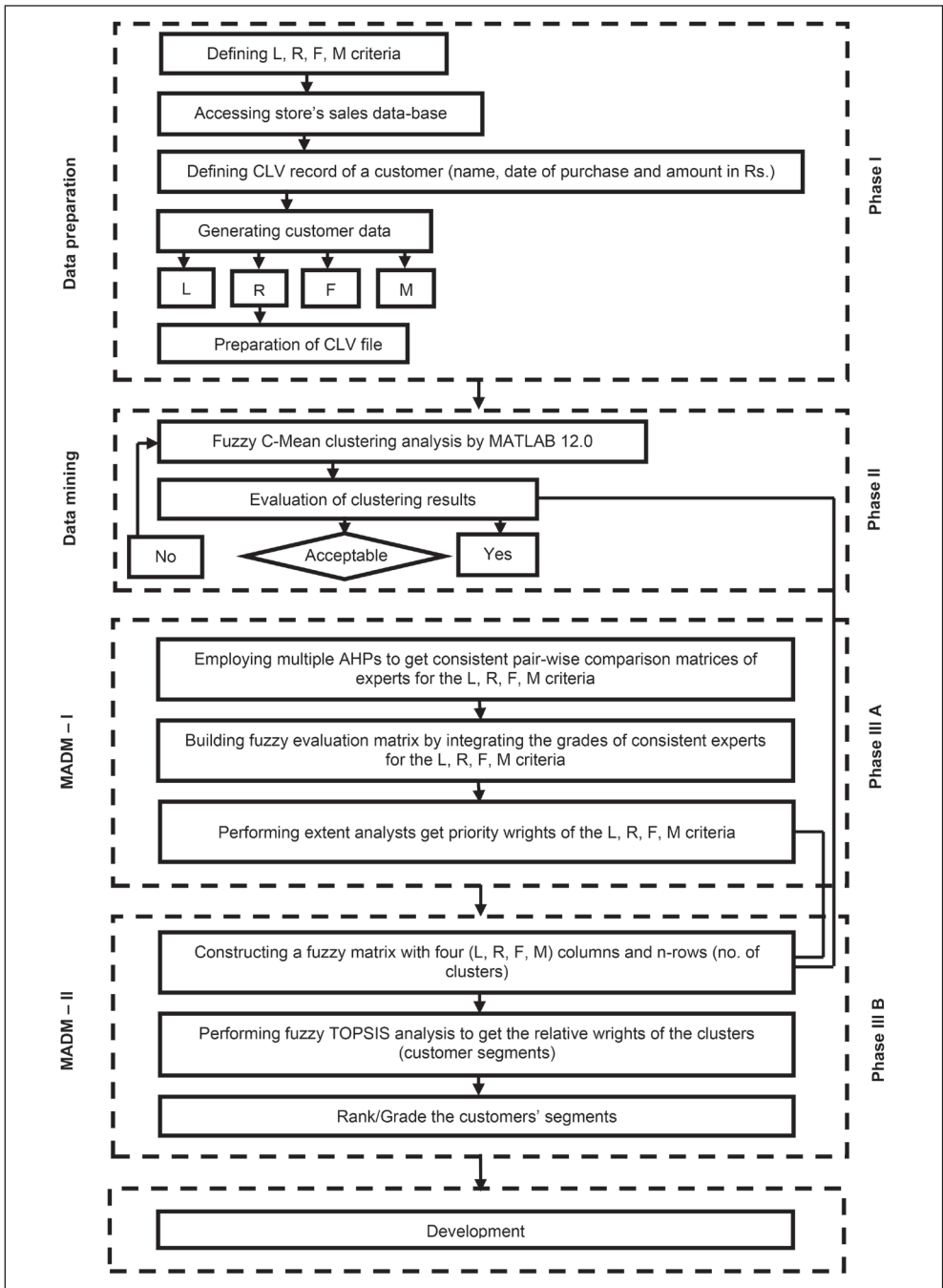
$$C_j = \frac{\sum_{i=1}^n U_{ij}^m \cdot x_i}{\sum_{i=1}^n U_{ij}^m}. \quad (3)$$

If  $\|U_{ij}^k - U_{ij}^{k-1}\| < \delta$ , then the iteration will be discontinued, where  $\delta$  is a prescribed accuracy level between 0 and 1, while  $k$  is the iteration step. This procedure converges to a local minimum or a saddle point of  $J_m$ . In our work, we use this method to cluster customers with similar LTV.

### 2.2 Fuzzy AHP (FAHP) for Ranking LRFM

Though Analytic AHP has been employed by many researchers for assigning weights to the variables L, R, F, M (Ray & Mangaraj, 2016), yet the proposed Fuzzy AHP method for the

Fig. 1: Model of fuzzy MADM in data mining for customer segmentation



Source: own

**Tab. 1: Relative degree of comparison for pairwise comparison of L, R, F and M variables**

Comparative importance	Description	Explanation
1	Equally important	Both the variable have equal contribution towards the objective
3	Weakly important	One variable is perceived slightly important in comparison to other
5	Strongly important	Importance of one variable with respect to goal is strong in comparison to other as per the expert's perception
7	Very strongly important	One variable is perceived to have very strong importance in comparison to other
9	Extremely important	One variable is perceived extremely important in comparison to other

The importance levels 2, 4, 6 and 8 represent the intermediate values. For example, 2 signifies intermediate between equally and weakly importance where judgment is between equally and slightly favoring one variable over another.

Source: own

same as the latter allows freedom to decision makers to provide their opinion through natural languages (Muduli & Barve, 2015; Shen et al., 2015).

This requires formulation of a fuzzy comparison matrix obtained from consistent comparison matrices from a set of AHPs depending upon the number of decision-makers (DMs). To carry out this, nine experts from sales department that include three experts from each of the three different management layers were selected and requested to provide their judgement using a 9 point scale described in Tab. 1, about how important they perceive one variable in comparison to the other. These scores were used to develop the pairwise comparison matrices.

### 2.3 Evaluation of Weights for LRFM variables

The steps followed to assess the weights of the LRFM variables  $w_L$ ,  $w_R$ ,  $w_F$  and  $w_M$  are as follows:

*Step 1: Development of pair-wise comparison matrix*

This asks the decision-makers to make pair-wise comparison of the LRFM variables using the scale as presented in Tab. 1. We show this comparison as a 4 x 4 judgment matrix for each DM. For example, the judgment matrix  $R_1$  for expert 1 with respect to the LRFM criteria could be presented as shown in formula (4).

$$R_1 = \begin{matrix} & \begin{matrix} L & R & F & M \end{matrix} \\ \begin{matrix} L \\ R \\ F \\ M \end{matrix} & \begin{bmatrix} r_{1,11} & r_{1,12} & r_{1,13} & r_{1,14} \\ r_{1,21} & r_{1,22} & r_{1,23} & r_{1,24} \\ r_{1,31} & r_{1,32} & r_{1,33} & r_{1,34} \\ r_{1,41} & r_{1,42} & r_{1,43} & r_{1,44} \end{bmatrix} \end{matrix} \quad (4)$$

In this matrix,  $r_{1,11} = r_{1,22} = r_{1,33} = r_{1,44}$  and  $r_{1,ij} \cdot r_{1,ji} = 1$ , where  $i = j = 1, 2, \dots, 4$  and  $r_{1,ij}$  is a value in the above scale where  $i \neq j$ .

*Step 2: Assess the consistency of pairwise comparison*

While making pairwise comparison, some DMs may make inconsistent judgments. This can be known by measuring the degree of inconsistency of all the DMs, where no inconsistency represents perfect consistency. As perfect consistency is rarely achieved in such a situation, an inconsistency index value of less than 0.1 is acceptable for a judgment matrix as a consistent. In this way, we get all the consistent judgment matrices for the consistent DMs.

*Step 3: Construct the fuzzy evaluation matrix*

We aggregate all these matrices to a fuzzy evaluation matrix where the elements of the matrix are the TFNs  $(l_{ijk}, m_{ijk}, u_{ijk})$  for  $k$  number of DMs. We determine the values of  $l_{ijk}$ ,  $m_{ijk}$ , and  $u_{ijk}$  using formulas (5), (6) and (7) respectively.

$$l_{ijk} = \min_k(j_{k,ij}), \quad (5)$$



$$m_{ijk} = \frac{\sum_k J_{k,ij}}{k}, \tag{6}$$

$$u_{ijk} = \max_k (J_{k,ij}). \tag{7}$$

For each value of  $i$  and  $j$ ,  $m_{ijk}$  gives the maximal grade of the membership function  $\mu_x$ , refer to formula (8), where  $l_{ijk}$  and  $u_{ijk}$  are the lower and upper bounds that limit the field of possible evaluation (Chang, 1996). Hence, for each  $i$  and  $j$ .

$$\mu_x = \begin{cases} \frac{(x-l_{ijk})}{(m_{ijk}-l_{ijk})} & x \in [l_{ijk}, m_{ijk}] \\ \frac{(u_{ijk}-x)}{(u_{ijk}-m_{ijk})} & x \in [m_{ijk}, u_{ijk}] \\ 0 & otherwise \end{cases} \tag{8}$$

Hence, we construct the fuzzy evaluation matrix for  $k$  number of DMs using formula (9).

$$F_E = \begin{matrix} & \begin{matrix} L & R \end{matrix} \\ \begin{matrix} L \\ R \\ F \\ M \end{matrix} & \begin{bmatrix} (l_{11k}, m_{11k}, u_{11k}) & (l_{12k}, m_{12k}, u_{12k}) \\ (l_{21k}, m_{21k}, u_{21k}) & (l_{22k}, m_{22k}, u_{22k}) \\ (l_{31k}, m_{31k}, u_{31k}) & (l_{32k}, m_{32k}, u_{32k}) \\ (l_{41k}, m_{41k}, u_{41k}) & (l_{42k}, m_{42k}, u_{42k}) \end{bmatrix} \\ \begin{matrix} F & M \end{matrix} & \begin{bmatrix} (l_{13k}, m_{13k}, u_{13k}) & (l_{14k}, m_{14k}, u_{14k}) \\ (l_{23k}, m_{23k}, u_{23k}) & (l_{24k}, m_{24k}, u_{24k}) \\ (l_{33k}, m_{33k}, u_{33k}) & (l_{34k}, m_{34k}, u_{34k}) \\ (l_{43k}, m_{43k}, u_{43k}) & (l_{44k}, m_{44k}, u_{44k}) \end{bmatrix} \end{matrix} \tag{9}$$

**Step 4: Obtain the weights of LRFM criteria**

Extent analysis for FAHP proposed by Chang (1996) was conducted to obtain normalized non-fuzzy weight of the LRFM variables  $w_L, w_R, w_F$  and  $w_M$ , whose sum is equal to one.

**2.4 Fuzzy TOPSIS (FTOPSIS) for Ranking Clusters**

Technique for Order Preference (TOPSIS) is also a popular MCDM technique and used to determine relative weights of a set of alternatives based on a finite number of decision criteria (Tian et al., 2018; Bharti, 2020; Juan et al., 2021; Liao et al., 2020) as shown in Fig. 2. However, unlike a comparison matrix of values for alternatives with respect to a criterion, this method is based on a matrix of actual values of alternatives for a set of weighted criteria. As gets the weights of LRFM variables by FAHP, look for actual values of the customer clusters having similar LTV with respect to LRFM criteria in order to get

the relative weights of these clusters. The most ideal option evaluation through FTOPSIS is based on the computation of shortest distance to the ideal solution (Wailoni et al., 2022). Hence, a positive ideal solution (PIS) needs to be calculated from the matrix as well as the weighted criteria. Similarly, a negative ideal solution (NIS) also computed. In this case, The PIS is a four-dimensional co-ordinate out of which two are the weighted maximum values of these clusters for the frequency and monetary value criteria, whereas the other two are the weighted minimum values of the same clusters for recency and length criteria. The best cluster is the one, which is closest to the PIS and at the same time, farthest from the NIS. In this work, we employ fuzzy TOPSIS (FTOPSIS) as the minimum, average and maximum values of LRFM variable for different clusters are obtained by FCM clustering approach and hence, are expressed as TFNs in the criteria-alternative matrix.

The steps followed to develop the FTOPSIS model are as follows:

1. Cluster of customers is obtained by FCM clustering approach.
2. Weight of each LRFM criterion is calculated employing FAHP.
3. Establishing the data matrix for the FTOPSIS as presented in Tab. 2.

In Tab. 2, each cell value represents a TFN. For example,  $(L_{il}, L_{im}, L_{iu})$  for  $i = 1, 2, \dots, s$  denotes TFN, where  $L_{il}, L_{im}$  and  $L_{iu}$  are the lower, average and upper values of the variable  $L$  obtained by FCM clustering analysis across the generated clusters. For each variable crisp values are obtained using formula (10).

$$\Phi_L = \begin{cases} \frac{(L-L_{il})}{(L_{im}-L_{il})} & y \in [L_{il}, L_{im}] \\ \frac{(L_{iu}-L)}{(L_{iu}-L_{im})} & y \in [L_{im}, L_{iu}] \\ 0 & otherwise \end{cases} \tag{10}$$

where  $i = 1, 2, \dots, s$ . Hence, we get maximal grade of the membership function of  $L$ , i.e. 1 at  $L_{im}$ . Similarly,  $L_{il}$  and  $L_{iu}$  are the points for zero satisfaction of the same function. As discussed in the previous sub-section, we denote  $w_L, w_R, w_F$  and  $w_M$  are the relative weights of the LRFM variables.

**Tab. 2: Data matrix for FTOPSIS**

Clusters	Length ( $w_L$ )	Recency ( $w_R$ )	Frequency ( $w_F$ )	Monetary value ( $w_M$ )
Cluster 1	( $L_{1l}, L_{1m}, L_{1u}$ )	( $R_{1l}, R_{1m}, R_{1u}$ )	( $F_{1l}, F_{1m}, F_{1u}$ )	( $M_{1l}, M_{1m}, M_{1u}$ )
Cluster 2	( $L_{2l}, L_{2m}, L_{2u}$ )	( $R_{2l}, R_{2m}, R_{2u}$ )	( $F_{2l}, F_{2m}, F_{2u}$ )	( $M_{2l}, M_{2m}, M_{2u}$ )
...	...	...	...	...
...	...	...	...	...
Cluster S	( $L_{sl}, L_{sm}, L_{su}$ )	( $R_{sl}, R_{sm}, R_{su}$ )	( $F_{sl}, F_{sm}, F_{su}$ )	( $M_{sl}, M_{sm}, M_{su}$ )

Source: own

4. We normalize the data matrix  $[(x_{ij}, x_{im}, x_{iu})]_{s \times 4}$ , where  $x = L, R, F$  and  $M$  and get it as:

$$D = [(r_{ij}, r_{im}, r_{iu})]_{s \times 4}$$

$$\text{where } r_{ij} = \frac{x_{ij}}{\sum_{i=1}^s x_{ij}}; r_{im} = \frac{x_{im}}{\sum_{i=1}^s x_{im}};$$

$$r_{iu} = \frac{x_{iu}}{\sum_{i=1}^s x_{iu}}$$

5. We compute the weighted matrix

$$V = [(v_{ij}, v_{im}, v_{iu})]_{s \times 4}$$

$$\text{where } v_{ij} = w_j \cdot r_{ij}; v_{im} = w_j \cdot r_{im}; v_{iu} = w_j \cdot r_{iu}$$

and  $i = 1, 2, 3, \dots, s$ .

6. We determine the positive ideal point  $A^+$  and the negative ideal point  $A^-$  as:

$$A^+ = (v_1^+, v_2^+, v_3^+, v_4^+) \text{ and } A^- = (v_1^-, v_2^-, v_3^-, v_4^-)$$

where  $v_j^+ = \max_i(v_{ij})$  and  $v_j^- = \min_i(v_{ij})$ :  
 $i = 1, 2, 3 \dots s$

7. We find out the distance of each cluster  $C_i = (v_{i1}, v_{i2}, v_{i3}, v_{i4}) : i = 1, 2, 3 \dots s$  from  $A^+$  and  $A^-$  as  $d_i^+$  and  $d_i^-$  respectively,

$$\text{where } d_i^+ = [\sum_{j=1}^4 (v_{ij} - v_j^+)^2]^{1/2} \text{ and}$$

$$d_i^- = [\sum_{j=1}^4 (v_{ij} - v_j^-)^2]^{1/2}; i = 1, 2, 3, \dots, s$$

8. We compute the relative closeness of each cluster from the ideal solutions as:

$$R_i = \frac{d_i^-}{d_i^+ + d_i^-}; i = 1, 2, 3, \dots, s$$

Hence, higher value of  $R_i$  implies better CLV score of  $i^{th}$  cluster.

### 3. A Case Study of a Retail Store

For the evaluation of the proposed model, a real-life case study of Indian retail firm is chosen. The firm has several stores located in different parts of the country i.e., India and has hundreds of regular customers. However, this work considered the retail store located at Bhubaneswar, Odisha, India. The reason to choose this particular store for experimental work is that it is close to the median value of the customers are concerned. After discussion with the firm management for an improved CRM,

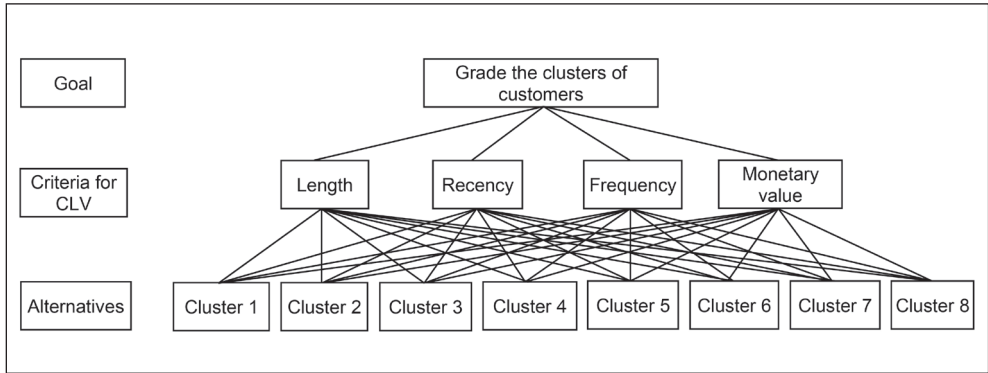
decided for implementation of the proposed modelling approach to its customers database for the segmenting of the customers in order to have different promotional measures, as an important component of its marketing strategy.

The store's sales database is designed having name, card number, date of purchase and amount of purchase as the attributes related to customers. Then out of that L, R, F and M values of a particular customers transactions in last 18 months are identified. The size of the customers database is 1,600, who visited the store multiple in last 18 months. For segmenting these customers into multiple groups along L, R, F and M values. The same values used in FCM clustering analysis by MATLAB 12.0. Obtained eight clusters with their corresponding minimum, average and maximum values of L, R, F and M criteria as presented in Tab. 3.

Based on the definition of L, R, F and M criteria for defining CLV as explained in section 2, this work evaluated the relative importance of each of them with respect to nine experts from the store. These experts included chief marketing manager, business manager, sales' manager, logistic manager and administrative manager of the store. They provided their individual pairwise comparison of the data of these criteria in the 9-point scale as described in Tab. 1. When assessed for their consistency, four of them were found to be consistent while the fifth one was asked to revise his data. After getting the consistent matrices, constructed the fuzzy evaluation matrix as discussed in step 3. Then employed extent analysis to handle this matrix to obtain the normalized non-fuzzy relative weights of the criteria as shown in Tab. 3. After getting the relative importance of the L, R, F and M criteria as well as their minimum, average and maximum values of eight clusters, constructed the fuzzy data matrix of these clusters with



**Fig. 2: Multi attribute decision making process**



Source: own

respect to the L, R, F and M criteria where each cell represented a value of a cluster expressed as triangular fuzzy number. Thus, organized the L, R, F and M values of these clusters as an 8 x 4 matrix for FTOPSIS analysis as discussed in section 2.3.

The decision variables in this problem have different units of measurement. Hence, to counter this issue a normalized data matrix was derived from Tab. 3 following step 4 discussed in section 2.4, and shown in Tab. 4.

The normalized data matrix is then converted to a weighted normalized data matrix as shown in Tab. 5 by using the weights of LRFM variables. Computational procedure is discussed in step 5 in section 2.4.

Following computational procedure discussed in steps 6 and step 7 of section 2.4, FNIS and FPIS are calculated from using Tab. 5 values and are shown in Tab. 6 and Tab. 7 respectively.

**Tab. 3: Clusters created by fuzzy c-means technique through MATLAB 12.0**

Cluster No.	Length (0.1120)			Recency (0.1971)			Frequency (0.2446)			Monetary value (0.4463)			No. of customers
	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value	
1	360	340	349	190	151	162	12	6	9	233,898	233,898	233,898	400
2	724	307	453	400	22	273	30	11	24	377,487	242,280	225,840	370
3	693	342	483	361	123	229	23	14	18	261,963	29,945	176,691	402
4	724	307	491	400	163	306	22	1	10	209,155	25,863	103,290	311
5	724	307	508	400	61	268	21	1	9	377,257	165,804	297,820	3
6	724	311	566	319	5	99	30	5	22	348,549	26,101	125,440	9
7	724	307	597	297	5	101	30	1	10	377,590	35,034	261,710	84
8	695	307	442	237	5	119	19	1	9	259,665	24,909	112,820	21
<b>Total</b>	<b>5,368</b>	<b>2,528</b>	<b>3,889</b>	<b>2,604</b>	<b>535</b>	<b>1,557</b>	<b>187</b>	<b>40</b>	<b>111</b>	<b>2,445,564</b>	<b>783,534</b>	<b>1,537,509</b>	<b>1,600</b>
<b>Average</b>			<b>486.125</b>			<b>194.625</b>			<b>13.875</b>			<b>192,188.625</b>	

Source: own

Tab. 4: Normalized data matrix

Cluster No.	Length (0.1120)			Recency (0.1971)			Frequency (0.2446)			Monetary value (0.4463)		
	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value
1	0.06706	0.13449	0.08974	0.07296	0.28224	0.10405	0.06417	0.15000	0.08108	0.09564	0.29840	0.15213
2	0.13487	0.12144	0.11648	0.15361	0.04112	0.17534	0.16043	0.27500	0.21622	0.15436	0.30910	0.14689
3	0.12910	0.13528	0.12420	0.13863	0.22991	0.14708	0.12299	0.35000	0.16216	0.10712	0.03820	0.11492
4	0.13487	0.12144	0.12625	0.15361	0.30467	0.19653	0.11765	0.02500	0.09009	0.08552	0.03300	0.06718
5	0.13487	0.12144	0.13062	0.15361	0.11402	0.17213	0.11230	0.02500	0.08108	0.15426	0.21153	0.19370
6	0.13487	0.12302	0.14554	0.12250	0.00935	0.06358	0.16043	0.12500	0.19820	0.14252	0.03330	0.08159
7	0.13487	0.12144	0.15351	0.11406	0.00935	0.06487	0.16043	0.02500	0.09009	0.15440	0.04470	0.17022
8	0.12947	0.12144	0.11365	0.09101	0.00935	0.07643	0.10160	0.02500	0.08108	0.10618	0.03178	0.07338

Source: own

Tab. 5: Weighted normalized data matrix

Cluster No.	Length (0.1120)			Recency (0.1971)			Frequency (0.2446)			Monetary value (0.4463)		
	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value	Highest value	Lowest value	Average value
1	0.00751	0.01506	0.01005	0.01438	0.05563	0.02051	0.01570	0.03669	0.01983	0.04268	0.13318	0.06789
2	0.01511	0.01360	0.01305	0.03028	0.00811	0.03456	0.03924	0.06727	0.05289	0.06889	0.13795	0.06556
3	0.01446	0.01515	0.01391	0.02732	0.04531	0.02899	0.03008	0.08561	0.03966	0.04781	0.01705	0.05129
4	0.01511	0.01360	0.01414	0.03028	0.06005	0.03874	0.02878	0.00612	0.02204	0.03817	0.01473	0.02998
5	0.01511	0.01360	0.01463	0.03028	0.02247	0.03393	0.02747	0.00612	0.01983	0.06885	0.09441	0.08645
6	0.01511	0.01378	0.01630	0.02415	0.00184	0.01253	0.03924	0.03058	0.04848	0.06361	0.01486	0.03641
7	0.01511	0.01360	0.01719	0.02248	0.00184	0.01279	0.03924	0.00612	0.02204	0.06891	0.01995	0.07597
8	0.01450	0.01360	0.01273	0.01794	0.00184	0.01506	0.02485	0.00612	0.01983	0.04739	0.01418	0.03275
Max	0.01511	0.01515	0.01719	0.03028	0.06005	0.03874	0.03924	0.08561	0.05289	0.06891	0.13795	0.08645
Min	0.00751	0.01360	0.01005	0.01438	0.00184	0.01253	0.01570	0.00612	0.01983	0.03817	0.01418	0.02998

Source: own

Tab. 6: Computation of distance between each criterion and corresponding +Ve ideal value for each alternative

Cluster No.	Distance for L	Distance for R	Distance for F	Distance for M	$d_i^+$
1	0.00084	0.03139	0.01765	0.07215	0.12204
2	0.00471	0.01609	0.04237	0.07644	0.13962
3	0.00468	0.02786	0.04803	0.01360	0.09416
4	0.00498	0.03798	0.00766	0.00031	0.05093
5	0.00512	0.01946	0.00680	0.05934	0.09072
6	0.00568	0.00564	0.02565	0.01515	0.05212
7	0.00602	0.00468	0.01365	0.03211	0.05646
8	0.00432	0.00252	0.00529	0.00556	0.01769

Source: own

**Tab. 7: Computation of distance between each criterion and corresponding –Ve ideal value for each alternative**

Cluster No.	Distance for L	Distance for R	Distance for F	Distance for M	$d_i^+$
1	0.00602	0.01419	0.03670	0.01875	0.07566
2	0.00256	0.03009	0.01059	0.01206	0.05530
3	0.00193	0.01034	0.00929	0.07371	0.09527
4	0.00198	0.00000	0.04960	0.08024	0.13182
5	0.00300	0.02187	0.05017	0.02514	0.10018
6	0.00095	0.06413	0.03188	0.07677	0.17372
7	0.00155	0.03707	0.04923	0.06840	0.15625
8	0.00275	0.03697	0.05040	0.07888	0.16900

Source: own

**Tab. 8: Computation of closeness coefficient and ranking of clusters**

Cluster No.	$d_i^-$	$d_i^+$	$c_i$	Ranks
1	0.07566	0.12204	0.61730	2
2	0.05530	0.13962	0.71629	1
3	0.09527	0.09416	0.49709	3
4	0.13182	0.05093	0.27869	5
5	0.10018	0.09072	0.47523	4
6	0.17372	0.05212	0.23077	7
7	0.15625	0.05646	0.26543	6
8	0.16900	0.01769	0.09473	8

Source: own

#### 4. Discussions and Implications

Owing to the rapid globalization a large number of competitive products are available near the customers. Detailed information about these products such as their specifications, contents, price, merits, demerits, alternate products available in the market etc, are also available with the customer due to the advancement of information communication technologies. Consequently, customers are evaluating the products according to the criteria which they consider to be more important for themselves. Consumer choices are also getting affected by social media campaigns. Hence, it became important for the organisations to shift their focus from brand and product to the customer (Mosaddegh et al., 2021). Companies are emphasizing on developing strategies that will

encourage long-term relationships with their clients instead of having discrete transactions (Mosaddegh et al., 2021; De Marco et al., 2021). This requires understanding the needs of each client, however as the resources are limited, many organizations prefer to identify the valuable customers and allocate maximum resources to them (Mosaddegh et al., 2021) which necessitates segmentation of the customers. In this regard, this research attempted to segment the customers of Indian retail sector employing Fuzzy C-Mean Clustering technique. This research categorized the customers into 8 clusters as shown in Tab. 3.

Analysis of the Tab. 3 values shows that average value of 'L' for all clusters is 486.125 and cluster 7, cluster 6, cluster 5 and cluster 4

have higher values of L than the average value. This indicates that the customers in these clusters are long term customers. Similarly, if the average value of variable 'R' of each cluster is compared with that of the average value for all the clusters, it is inferred that customers of cluster 4, cluster 5, cluster 2 and cluster 3 are uncertain or lost customers. Again if the average values of 'F' for each cluster is compared with that of the total average, it could be seen that customers in cluster 2, cluster 6 and cluster 3 are more frequent. Further, the comparison based on consumption value (Variable M) indicates that cluster 5, cluster 7, cluster 1 and cluster 2 have customers which are more important for the company in terms of their consumption.

Few previous studies evaluated CLV score, using LRFM while giving equal weightage to all the variables (Rahmadiani et al., 2020) while few studies employed weighted approach and employed AHP for this purpose (Rad et al., 2011; Ray & Mangaraj, 2016). The LRFM variables have different measurement units hence provision of equal weightage is not justified. Again, use of AHP for computation of AHP fails to capture the vagueness associated with human judgement. To counter these facts, Fuzzy AHP employed here to allocate the weights to the LRFM variables followed by Fuzzy TOPSIS to rank these clusters. Tab. 8 shows the ranking order of the alternatives with the Fuzzy TOPSIS technique which is as follows: cluster 2 > cluster 1 > cluster 3 > cluster 5 > cluster 4 > cluster 7 > cluster 6 > cluster 8.

The ranking indicates that the customers placed in cluster 2 are most valuable for the organizations. Hence, maximum resources need to be allocated to satisfy these customers. This research also identified cluster 8 contains the least valuable customers. Hence, special strategy needs to be formulated to retain these customers. Segmentation of the customers also facilitates developing different loyalty program strategies for different segments to enhance the effectiveness of these programs.

## Conclusions

This paper proposed an integrated Fuzzy AHP-Fuzzy TOPSIS based framework to rank the Cluster of Customers based on their CLV score. Four decision criteria L (length), R (recency), F (frequency) and M (monetary value) are considered in this research to evaluate CLV

score using the integrated framework. These criteria were evaluated to determine the order of Cluster alternatives for selecting the most appropriate one. Experts having adequate experience in sales department of Indian retail firms were asked to judge the relative importance of the four criteria in terms of linguistic variables. The opinion of the experts were utilized to develop fuzzy comparison matrix. The fuzzy comparison matrix is further processed to assess closeness co-efficient of each alternatives (customer cluster in our research) with the ideal value. Ranking of the customer clusters in this research is based on the closeness coefficient of the clusters with the ideal one. This study revealed that cluster 2 is the best alternative while cluster-8 is the least important alternative for having highest and lowest CLV values respectively. The order of the cluster is as follows: 2 > 1 > 3 > 5 > 4 > 7 > 6 > 8. The findings of this research can assist organizations to make long-term relationship with customers by formulating appropriate strategies. The top management of the Retail Company can formulate different cluster specific suitable strategies for the further growth of the business. Future research could be focused on conducting association analysis to establish customer buying patterns to explore which products have been purchased together frequently by which customers or customer groups; enhancing the merchant's web site to enable a consumer's shopping activities to be captured and tracked instantaneously and accurately; and predicting each customer's lifecycle value to quantify the level of diversity of each customer.

## References

- Aghdaie, M. H., Hashemkhani Zolfani, S., & Zavadskas, E. K. (2014). Synergies of data mining and multiple attribute decision making. *Procedia Social and Behavioural Sciences*, 110, 767–776. <https://doi.org/10.1016/j.sbspro.2013.12.921>
- Amiri, M., Hashemi-Tabatabaei, M., Ghahremanloo, M., Keshavarz-Ghorabae, M., Zavadskas, E. K., & Banaitis, A. (2021). A new fuzzy BWM approach for evaluating and selecting a sustainable supplier in supply chain management. *International Journal of Sustainable Development and World Ecology*, 28(2), 125–142. <https://doi.org/10.1080/13504509.2020.1793424>

Anitha, P., & Patil, M. M. (2019). RFM model for customer purchase behavior using K-Means algorithm. *Journal of King Saud University – Computer and Information Sciences* (In press). <https://doi.org/10.1016/j.jksuci.2019.12.011>

Arabameri, A., Rezaei, K., Cerda, A., Lombardo, L., & Rodrigo-Comino, J. (2019). GIS-based groundwater potential mapping in Shahroud plain, Iran. A comparison among statistical (bivariate and multivariate), data mining and MCDM approaches. *Science of the Total Environment*, 658, 160–177. <https://doi.org/10.1016/j.scitotenv.2018.12.115>

Benoit, D. F., & Van den Poel, D. (2009). Benefits of quantile regression for the analysis of customer life time value in a contractual setting: An application in financial services. *Expert Systems with Applications*, 36, 10475–10484. <https://doi.org/10.1016/j.eswa.2009.01.031>

Bezdek, J. C., Ehrlich, R., & Full, W. (1984). FCM: The fuzzy c-means clustering algorithm. *Computers and Geosciences*, 10(2–3), 191–203. [https://doi.org/10.1016/0098-3004\(84\)90020-7](https://doi.org/10.1016/0098-3004(84)90020-7)

Bharti, P. S. (2020). Two-step optimization of electric discharge machining using neural network based approach and TOPSIS. *Journal of Interdisciplinary Mathematics*, 23(1), 81–96. <https://doi.org/10.1080/09720502.2020.1741222>

Chang, D.-Y. (1996). Applications of the extent analysis method on fuzzy AHP. *European Journal of Operational Research*, 95(3), 649–655. [https://doi.org/10.1016/0377-2217\(95\)00300-2](https://doi.org/10.1016/0377-2217(95)00300-2)

Chang, H. H., & Tsay, S. F. (2004). Integrating of SOM and K-mean in data mining clustering: An empirical study of CRM and profitability evaluation. *Journal of Information Management*, 11, 161–203.

Chiang, L.-L. L., & Yang, C.-S. (2018). Does country-of-origin brand personality generate retail customer lifetime value? A Big Data analytics approach. *Technological Forecasting and Social Change*, 130, 177–187. <https://doi.org/10.1016/j.techfore.2017.06.034>

Corne, D., Dhaenens, C., & Jourdan, L. (2012). Synergies between operations research and data mining: The emerging use of multi-objective approaches. *European Journal of Operational Research*, 221(3), 469–479. <https://doi.org/10.1016/j.ejor.2012.03.039>

De Marco, M., Fantozzi, P., Fornaro, C., Laura, L., & Miloso, A. (2021). Cognitive analytics management of the customer lifetime value: An artificial neural network approach.

*Journal of Enterprise Information Management*, 34(2), 679–696. <https://doi.org/10.1108/JEIM-01-2020-0029>

Dunn, J. C. (1973). A fuzzy relative of the ISODATA process and its use in detecting compact well-separated clusters. *Journal of Cybernetics*, 3(3), 32–57. <https://doi.org/10.1080/01969727308546046>

Egemen, M. (2021). A framework for buyers' house selection criteria vs. post-occupancy residential satisfaction levels in North Cyprus. *International Journal of Strategic Property Management*, 25(1), 50–64. <https://doi.org/10.3846/ijspm.2020.13725>

Guerard, J. B., Xu, G., & Markowitz, H. (2021). A further analysis of robust regression modeling and data mining corrections testing in global stocks. *Annals of Operations Research*, 303(1), 175–195. <https://doi.org/10.1007/s10479-020-03521-y>

Guo, J.-Q., Chiang, S.-H., Liu, M., Yang, C.-C., & Guo, K.-Y. (2020). Can machine learning algorithms associated with text mining from internet data improve housing price prediction performance? *International Journal of Strategic Property Management*, 24(5), 300–312. <https://doi.org/10.3846/ijspm.2020.12742>

Hughes, A. M. (1994). *Strategic database marketing*. Chicago, IL: Probus Publishing.

Juan, Y.-K., Hsu, Y.-C., & Chang, Y.-P. (2021). Site selection assessment of vacant campus space transforming into daily care centers for the aged. *International Journal of Strategic Property Management*, 25(1), 34–49. <https://doi.org/10.3846/ijspm.2020.13800>

Khajvand, M., Zolfaghar, K., Ashoori, S., & Alizadeh, S. (2011). Estimating customer lifetime value based on RFM analysis of customer purchase behavior: Case study. *Procedia Computer Science*, 3, 57–63. <https://doi.org/10.1016/j.procs.2010.12.011>

Kotler, P. (2003). *Marketing for hospitality and tourism* (5th ed.). Chennai: Pearson Education India.

Liao, H., Tang, M., Li, Z., & Lev, B. (2019). Bibliometric analysis for highly cited papers in operations research and management science from 2008 to 2017 based on essential science indicators. *Omega*, 88, 223–236. <https://doi.org/10.1016/j.omega.2018.11.005>

Liao, H. C., Ren, R. X., Antucheviciene, J., Šaparauskas, J., & Al-Barakati, A. (2020). Sustainable construction supplier selection by a multiple criteria decision-making method with

hesitant linguistic information. *E&M Economics and Management*, 23(4), 119–136. <https://doi.org/10.15240/tul/001/2020-4-008>

Liu, D.-R., & Shih, Y.-Y. (2005). Integrating AHP and datamining for product recommendation based on customer lifetime value. *Information & Management*, 42(3), 387–400. <https://doi.org/10.1016/j.im.2004.01.008>

Liou, J. J. H., Chang, M. H., Lo, H. W., & Hsu, M. H. (2021). Application of an MCDM model with data mining techniques for green supplier evaluation and selection. *Applied Soft Computing*, 109, 107534. <https://doi.org/10.1016/j.asoc.2021.107534>

Mahdiraji, H. A., Zavadskas, E. K., Kazeminia, A., & Abbasi Kamardi, A. (2019). Marketing strategies evaluation based on big data analysis: A CLUSTERING-MCDM approach. *Economic Research – Ekonomska Istraživanja*, 32(1), 2882–2898. <https://doi.org/10.1080/1331677X.2019.1658534>

Malik, M. M., Abdallah, S., & Ala'raj, M. (2018). Data mining and predictive analytics applications for the delivery of healthcare services: a systematic literature review. *Annals of Operations Research*, 270(1), 287–312. <https://doi.org/10.1007/s10479-016-2393-z>

Marques, F. C., Ferreira, F. A. F., Zopounidis, C., & Banaitis, A. (2020). A system dynamics-based approach to determinants of family business growth. *Annals of Operations Research*. <https://doi.org/10.1007/s10479-020-03524-9>

Meisel, S., & Mattfeld, D. (2010). Synergies of operations research and data mining. *European Journal of Operational Research*, 206(1), 1–10. <https://doi.org/10.1016/j.ejor.2009.10.017>

Moktadir, M. A., Mahmud, Y., Banaitis, A., Sarder, T., & Khan, M. R. (2021). Key performance indicators for adopting sustainability practices in footwear supply chains. *E&M Economics and Management*, 24(1), 197–213. <https://doi.org/10.15240/tul/001/2021-1-013>

Mohandes, S. R., Sadeghi, H., Mahdiyar, A., Durdjev, S., Banaitis, A., Yahya, K., & Ismail, S. (2020). Assessing construction labours' safety level: A fuzzy MCDM approach. *Journal of Civil Engineering and Management*, 26(2), 175–188. <https://doi.org/10.3846/jcem.2020.11926>

Mosaddegh, A., Albadvi, A., Sepehri, M. M., & Teimourpour, B. (2021). Dynamics of customer segments: A predictor of customer lifetime value. *Expert Systems with Applications*, 172, 114606. <https://doi.org/10.1016/j.eswa.2021.114606>

Muduli, K., & Barve, A. (2015). Analysis of critical activities for GSCM implementation in mining supply chains in India using fuzzy analytical hierarchy process. *International Journal of Business Excellence*, 8(6), 767–797. <https://doi.org/10.1504/IJBEX.2015.072309>

Olafsson, S., Li, X., & Wu, S. (2008). Operations research and data mining. *European Journal of Operational Research*, 187(3), 1429–1448. <https://doi.org/10.1016/j.ejor.2006.09.023>

Ozkaya, G., Timor, M., & Erdin, C. (2021). Science, Technology and Innovation Policy Indicators and Comparisons of Countries through a Hybrid Model of Data Mining and MCDM Methods. *Sustainability*, 13(2), 694. <https://doi.org/10.3390/su13020694>

Pérez-Martin, A., Pérez-Torregrosa, A., & Vaca, M. (2018). Big Data techniques to measure credit banking risk in home equity loans. *Journal of Business Research*, 89, 448–454. <https://doi.org/10.1016/j.jbusres.2018.02.008>

Pineda, P. J. G., Liou, J. J., Hsu, C. C., & Chuang, Y. C. (2018). An integrated MCDM model for improving airline operational and financial performance. *Journal of Air Transport Management*, 68, 103–117. <https://doi.org/10.1016/j.jairtraman.2017.06.003>

Rad, A., Naderi, B., & Soltani, M. (2011). Clustering and ranking university majors using data mining and AHP algorithms: A case study in Iran. *Expert Systems with Applications*, 38(1), 755–763. <https://doi.org/10.1016/j.eswa.2010.07.029>

Rahmadianti, R., Dhini, A., & Laoh, E. (2020, November). Estimating customer lifetime value using LRFM model in pharmaceutical and medical device distribution company. In *Proceedings of the 2020 International Conference on ICT for Smart Society (ICISS)* (pp. 1–5). Bandung, Indonesia. <https://doi.org/10.1109/ICISS50791.2020.9307592>

Ray, M., & Mangaraj, B. K. (2016). AHP Based Data Mining for customer segmentation based on customer lifetime value. *International Journal of Data Mining Techniques and Applications*, 5(1), 28–34. <https://doi.org/10.20894/IJDMTA.102.005.001.007>

Safari, F., Safari, N., & Montazer, G. A. (2016). Customer lifetime value determination based on RFM model. *Marketing Intelligence & Planning*, 34(4), 446–461. <https://doi.org/10.1108/MIP-03-2015-0060>

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(7), 5259–5264. <https://doi.org/10.1016/j.eswa.2009.12.070>

Shen, L., Muduli, K., & Barve, A. (2015). Developing a sustainable development framework in the context of mining industries: AHP approach. *Resources Policy*, 46(1), 15–26. <https://doi.org/10.1016/j.resourpol.2013.10.006>

Tian, M.-W., Liao, X.-Z., Wu, L.-X., Peng, L.-H., & Yan, S.-R. (2018). Multi-attribute decision making of target enterprises in mergers and acquisitions. *Journal of Interdisciplinary Mathematics*, 21(5), 1103–1108. <https://doi.org/10.1080/09720502.2018.1493037>

Tsai, H.-H. (2012). Global data mining: An empirical study of current trends, future forecasts and technology diffusions. *Expert Systems with Applications*, 39(9), 8172–8181. <https://doi.org/10.1016/j.eswa.2012.01.150>

Wailoni, X., Swain, S., Lafanama, S., & Muduli, K. (2022). Analytical approach for prioritizing waste management practices: Implications for sustainable development

exercises in health care sector. *International Journal of Social Ecology and Sustainable Development*, 13(1), 43. <https://doi.org/10.4018/IJSESD.289643>.

Wei, J.-T., Lin, S.-Y., Weng, C.-C., & Wu, H.-H. (2012). A case study of applying LRFM model in market segmentation of a children's dental clinic. *Expert Systems with Applications*, 39(5), 5529–5533. <https://doi.org/10.1016/j.eswa.2011.11.066>

Wiesel, T., Skiera, B., & Villanueva, J. (2011). Customer Lifetime Value and Customer Equity Models Using Company-reported Summary Data. *Journal of Interactive Marketing*, 25(1), 20–22. <https://doi.org/10.1016/j.intmar.2010.12.003>

Wu, D., & Olson, D. L. (2006). A TOPSIS Data Mining Demonstration and Application to Credit Scoring. *International Journal of Data Warehousing and Mining*, 2(3), 16–26. <https://doi.org/10.4018/jdwm.2006070102>

Zare, H., & Emadi, S. (2020). Determination of Customer Satisfaction using Improved K-means algorithm. *Soft Computing*, 24(11), 16947–16965. <https://doi.org/10.1007/s00500-020-04988-4>