



Contents lists available at FOMJ

# Fuzzy Optimization and Modelling Journal

Journal homepage: <https://sanad.iau.ir/journal/fomj/>**Paper Type: Research Paper**

## Customer Clustering via Combined TOPSIS-Fuzzy -K-MEANS Method to Design an Efficient Customer Relationship System: A Real-life Case Study in the Copper Industry

**Hossein Mohammadi Dolat-Abadi<sup>a,\*</sup>, Amir Sadra Sadat<sup>b</sup>**<sup>a</sup> Farabi College, Department of Industrial Engineering, University of Tehran, Iran<sup>b</sup> Farabi College, Department of Industrial Engineering, University of Tehran, Iran

### ARTICLE INFO

**Article history:**

Received 20 December 2023

Revised 23 February 2024

Accepted 13 March 2024

Available online 26 April 2024

**Keywords:**

K-means Method

Fuzzy TOPSIS

Customer Eelationship System

Metals Supply Chain

### ABSTRACT

The contemporary application of analytical approaches, including clustering, classification, and ranking, in customer analysis empowers supply chain members to effectively align their organizational and commercial objectives. This study introduces a clustering model designed to scrutinize customers within a metal supply chain, defining optimal strategies tailored to each cluster. These strategies contribute to the implementation of a comprehensive customer relationship system, fostering competitiveness in the market. To achieve this goal, the initial step involves the review, cleaning, and normalization of the company's customer data. These data comprise scores in eleven criteria aspects for each customer, encompassing aspects such as good account status, absence of bounced checks, timely payment, legal status, presence of personal or governmental support, reputation, brand value, internal business managers' comments, each customer's share of total purchases, and production capacity. Expert-derived weights are assigned to these criteria. Subsequently, the k-means clustering technique is employed and validated through the silhouette score. Post clustering, the Fuzzy TOPSIS method is utilized to rank the clusters, determining their respective positions. Finally, strategies and approaches for each cluster are formulated, considering factors such as monetary credit allocation, discount rates, and trust levels in product sales. Overall, this research pioneers a comprehensive framework that goes beyond traditional models, offering a strategic roadmap for supply chain members to navigate a competitive market, standardize communication, and foster long-term relationships with customers.

<sup>\*</sup> Corresponding authorE-mail address: [hmohammadi@ut.ac.ir](mailto:hmohammadi@ut.ac.ir) (Hossein Mohammadi Dolat-Abadi)

DOI:

## 1. Introduction

Clustering problems have various applications, such as data mining, data compression, and pattern recognition and classification. The idea of what constitutes a good group depends very much on the application, and there are many ways to find groups that meet different criteria. For example, in the research literature, approaches based on division and integration, such as random approaches such as CLARA [9], CLARANS [17], and methods based on neural networks [4] have been proposed.

Moreover, the escalating competition in today's metal market underscores the growing significance of customer value. This heightened value not only contributes to securing a larger market share but is also instrumental in establishing robust customer relationships and delivering competitive products. The supply chain, functioning as a comprehensive network spanning raw material conversion to final product production and associated information systems, plays a pivotal role [8]. Effective management of material and information flow, both upstream and downstream within the network, is crucial for optimizing supply chain performance and ensuring customer satisfaction. Moreover, customer engagement with the organization is facilitated through interactive behaviors arising from meaningful categorization [19]. It is evident that the creation of value, coupled with the cultivation of interactive and close relationships with customers, serves as a formidable competitive advantage in the dynamic market landscape [14].

Hence, it is evident that despite existing conditions such as maximum credit limits, discounts, maximum credit to tonnage, and the traditional trading structure in the Iranian copper industry, there is an urgent need for the development of customer-focused policies. The metal and copper sector in Iran is undergoing substantial expansion propelled by factors like cost-effective raw material acquisition, robust earnings from exports, relentless demand, and the absence of competitive pricing aligned with quality and increased production. The advancement in metal and steel industries has now reached a juncture where private sector steel production supports around 600 downstream industries. In addition to these notable challenges, it is crucial to acknowledge that this industry has the potential to emerge as a driving force in the country's economy, offering extensive employment opportunities. As the copper industry stands among the fastest-growing sectors in Iran, its growth significantly impacts national economic development, promising an increased prevalence of this vital industry [22].

The process of categorizing and clustering customers assumes paramount importance in deciphering intricate customer behavior patterns, affording organizations the ability to anticipate emerging trends, formulate strategic decisions, optimize profits, and align policies with overarching company objectives [6]. Rooted in the acknowledgment of finite organizational resources, both classification and clustering methodologies emerge as indispensable tools for precision-focused customer service [29]. This method entails the meticulous clustering and grouping of customers based on meticulously collected data, facilitating nuanced analysis of data groups exhibiting pronounced similarities. Through this classification modality, the voluminous data or customer pool undergoes reduction, finding placement into discrete classes characterized by shared specific attributes within each class. [15].

This study aims to pinpoint the most effective indicators for customer cluster analysis in the metal supply chain using the K-means and Fuzzy TOPSIS methods where such a combination has not been proposed in the existing literature. In this paper, a new method is proposed to choose the number of clusters for the K-means algorithm. The proposed method can suggest multiple values of K-means to users for cases in which different clustering results are obtained with different levels of detail required. Moving forward, we rank our clusters using Fuzzy TOPSIS to handle the inherent uncertainty in fuzzy hypotheses by translating them into defined verbal expressions. After ranking the clusters with the defined criteria, a customer relationship system in each cluster is presented based on the 4 key factors of monetary credit allocation, discount amounts, the level of trust in selling products for each customer, and transportation cost. This system helps to deal with customers in a systematic and integrated manner and to minimize individual tastes in communication with customers.

The remainder of the study is as follows. Section 2 depicts the related literature by presenting a categorized table. The proposed models besides data collection are presented in Sections 3 and 4. Section 5 reports the results using a case study in the metal supply chain. In the end, Section 6 discusses the conclusion of the paper and proposes future directions for the study.

## 2. Literature Review

Recent years have witnessed a surge in studies within the realm of customer clustering, each with diverse objectives. Some delve into credit clustering, while others strive for deeper comprehension, seeking to allocate distinct policies for enhanced customer management. In the upcoming section, we will delve into a comprehensive research review centered on clustering methods and the evolution of fuzzy methods (Table 1).

**Table 1. Related works on clustering methods under the MCDM approach**

Reference	Research Methodology	Scope/Target
[28]	K-means and AHP	Customer clustering for demand forecasting
[21]	AHP and K-means	Research on customer classification by clustering
[27]	A fuzzy-based algorithm with a hierarchical analysis structure	To group the customers into multiple clusters
[1]	RFM, SOM, Neural network	Customer Credit Clustering for Presenting Appropriate Facilities
[3]	Fuzzy c-means GA	Customer clustering
[25]	K-means and Elbow Method	Identification of The Best Customer Profile Cluster
[13]	BWM, COPRAS, RFM	Proposing a digital banking strategy to bank customers
[2]	K-means and Vikor	grouping the house situation
[11]	Fuzzy AHP and K-means	analyzing the barriers to the adoption of Industry 4.0 practices
[24]	K-means and PF-VIKOR	Clustering failure modes
[5]	Fuzzy AHP-ARAS and K-means	Clustering for logistics hub location
[10]	AHP, K-means, Kohonen neural network	workshops clustering
[20]	K- means and TOPSIS	The best clustering results are ranked to produce alternative decisions in cluster selection
[23]	K-means and Fuzzy AHP	Specify climatic zones.
[26]	Fuzzy C-means and BWM	Prioritizing health, safety, and environmental risks
[18]	TOPSIS and K-means	Evaluating and Ranking Approach for Banks
[7]	K- means and ANN	Clustering Wireless Sensor Networks
The current paper	K-means and Fuzzy TOPSIS	Customer Clustering to Design a Customer Relationship System in the Metal Company

According to Table 1, it can be seen that the combination of multi-criteria decision-making methods with K-means has been significantly worked. However, the combination of K-means and Fuzzy TOPSIS has not been investigated, which this paper shed light on. In addition, according to Table 1, clustering in the fields of telecommunications, banking, environment, etc. has been investigated while the application of clustering in the supply chain of metals has not been reported, this is what we will address in this paper. Aligned with the examination of previous studies, it becomes apparent that a predominant number of research initiatives have leaned towards employing RFM and LRFM methods for customer clustering. Within this landscape, certain studies have integrated Multiple Criteria Decision-Making (MCDM) techniques like TOPSIS, while others have opted for data analysis methods such as SOM. Recognizing the existing gaps in this research domain, our study strategically diverges by embracing the K-means method coupled with Fuzzy TOPSIS.

### 3. Methodology

The K-means method serves as our clustering tool, strategically grouping observations into K clusters by assigning each observation to the cluster with the closest average, where this average acts as a reference sample. Following this clustering phase, our next step involves ranking these clusters through the application of the fuzzy TOPSIS technique. Now, the TOPSIS method is a gem among multiple criteria decision-making (MADM) methods, specifically designed for ranking options. It hinges on two crucial concepts: the "ideal solution" and "similarity to the ideal solution." The ideal solution, true to its name, embodies the best possible solution across all aspects, although achieving it in practice is often unattainable. Our goal is to get as close to this ideal as possible. To gauge how similar a plan (or option) is to the ideal and anti-ideal solutions, we calculate the distance between the plan (or option) and both these solutions. Subsequently, options are ranked based on the ratio of their distance from the anti-ideal solution. The ideal solution's evaluation and ranking, on the other hand, take into account the total distance from both the ideal and anti-ideal solutions.

#### 3.1. Clustering with K-means Clustering

The K-means algorithm operates through iterative processes, aiming to categorize a given dataset into non-overlapping and distinct subgroups known as clusters. Each data point, representing a record, is exclusively assigned to a single group within these clusters. The primary objective of the algorithm is to maximize the similarity of data points within a cluster while simultaneously maximizing the dissimilarity (distance) between clusters. This process involves dividing the data into clusters in a way that minimizes the sum of squared distances between data points and the center of their respective clusters. The algorithm seeks to achieve minimal diversity within clusters, resulting in greater homogeneity or similarity among the records within each cluster:

1. In this algorithm, K is the number of clusters that must be determined first.
2. According to the value of K, the points in our data are selected as the initial centers of the clusters. It should be noted that in this study these points were chosen randomly.
3. Using the Euclidean distance formula, the distance of each existing data with the centers of the previous step is obtained. We assign each data to the closest cluster.

$$d(x, y) = |x - y| = \sqrt{\sum_{i=1}^n d(x_i - y_i)^2} \quad (1)$$

Calculating the average of all data at all levels of a cluster and moving the new center of the cluster there.

4. Steps 3 to 5 continue until no data points from one cluster are assigned to another cluster.
5. Calculations are done by considering the weights of criteria in clustering.

The silhouette score functions as a pivotal metric for assessing the efficacy of a clustering algorithm. Leveraging both within-cluster distance (intra-cluster distance) and separation between clusters (inter-cluster distance), this metric computes a comprehensive score reflective of our clustering algorithm's performance. The silhouette score essentially measures the likeness of an object to its cluster, comparing it against the separation from other clusters. This score is on a scale from -1 to +1, where a high value indicates close alignment of the object with entities within its cluster and weak connections to neighboring clusters. In essence, when the majority of objects boast high silhouette scores, it suggests an apt cluster configuration. On the flip side, if a significant portion of points yield low or negative scores, it may imply an inadequate configuration with either too many or too few clusters. The steps of this method are described below:

1. Calculate the average distance of object i with all other objects in the same cluster, denoting this value as A(i).
2. Calculate the average distance of object i with all objects from different clusters and select the lowest average distance, referred to as B(i).
3. After obtaining these two values mentioned above, the silhouette coefficient is computed using the following formula:

$$SC(i) = \frac{B(i)-A(i)}{(A(i),B(i))}. \tag{2}$$

### 3.3. Classification of Clusters using Fuzzy TOPSIS

Much like the ideal solution in TOPSIS, the priority order technique operates on the principle that the chosen alternative should minimize the geometric distance from the Positive Ideal Solution (PIS) while maximizing the geometric distance from the anti-ideal solution (NIS). The procedural steps align with conventional TOPSIS methodology, with the key distinction being the incorporation of triangular fuzzy numbers in the calculations for this study. The following discusses the summarized steps of fuzzy TOPSIS:

1. Create a decision matrix based on options and evaluation criteria. If the fuzzy numbers are triangular:

$$r_{ij} = \left( \frac{a_{ij}}{c_j^*}, \frac{b_{ij}}{c_j^*}, \frac{c_{ij}}{c_j^*} \right) \tag{3}$$

where  $c_j^*$  is the maximum value in criterion  $j$  among all alternatives. If the fuzzy numbers are not triangular:

$$r_{ij} = \left( \frac{a_j^0}{c_{ij}}, \frac{a_j^0}{b_{ij}}, \frac{a_j^0}{c_{ij}} \right) \tag{4}$$

where  $a_j^0$  is the minimum value of an in criterion  $j$  among all options.

2. Multiply the weight matrix by the scale-free fuzzy matrix to derive the weighted fuzzy matrix. Determine the fuzzy ideal for each criterion:

$$A^+ = (V_1^+, V_2^+, \dots, V_N^+) \tag{5}$$

$$A^- = (V_1^-, V_2^-, \dots, V_N^-) \tag{6}$$

3. Calculate the fuzzy distance of each option to each criterion and obtain the sum of these values for each option. Fuzzy distance relationship for two numbers  $X$  and  $Y$ :

$$X = (x_1, x_2, x_3) \tag{7}$$

$$Y = (y_1, y_2, y_3) \tag{8}$$

$$D = \sqrt{1/3[(y_1 - x_1)^2 + (y_2 - x_2)^2 + (y_3 - x_3)^2]} \tag{9}$$

4. Then, calculate the sum of this distance for each option.

$$d_i^+ = \sum_{j=1}^n d[V_{ij} - V_{ij}^*] \tag{10}$$

$$d_i^- = \sum_{j=1}^n d[V_{ij} - V_{ij}^0] \tag{11}$$

where  $V_{ij}^*$  and  $V_{ij}^0$  are positive and negative ideals.

5. In the final step, calculate the Comprehensive Credit (CC):

$$CC_I = \frac{d_i^-}{d_i^+ + d_i^-} \tag{12}$$

## 4. Data Collection

This study taps into data sourced from a copper-based alloy producer embedded within the metals supply chain. The dataset encompasses comprehensive information concerning the organization's customers across multiple criteria. As a result, the statistical population under scrutiny comprises the customers affiliated with a copper-based alloy producer located in Tehran province. The research data, extracted from the company's database, precisely targets information about customers from the year 2022. Ultimately, the analysis focuses on data from 50 customers, assessed against 11 distinct criteria. These criteria have been extracted through interviews with managers and shareholders of the company. In the process of extracting the criteria, it has been tried to make practical data available about the criteria. It is important to note that the weights assigned to these criteria were determined through an expert-based method where the weights are not fixed but can take any value from the given intervals, so the score of each alternative is the maximum value that the weighted average can reach when the weights belong to those intervals. [12]. These weights were obtained directly from experts within the organization. In Table 2, you can find details about the customers and their scores in the respective criteria (the notation "N" indicates NULL values). Each criterion includes an interval range of weights, the minimum of that range is the lowest and the ceiling of that range is the most important. Each customer in each

criterion, according to their performance and experts' opinions, obtains a number in the defined range that indicates the weight of the customer in each criterion.

**Table 2. Customers and Points in the Criteria**

Customer	Good Banking record (1-10)	Bounced check (1-15)	Timely payment (1-10)	real/legal (1-5)	Support (1-10)	Good name (1-10)	Brand (1-5)	Expert opinions of the business manager (1-5)	Expert opinions of the manager experts (1-5)	Share of purchase With the total purchase (1-10)	Production and consumption capacity (1-15)
C1	10	15	10	5	10	10	5	5	5	2	15
C2	10	15	10	5	10	10	5	5	5	3	15
C3	10	15	10	5	10	10	5	5	5	5	15
C4	7	15	8	5	10	10	5	5	4	6	15
C5	8	12	10	5	10	10	5	5	5	9	10
C6	10	15	10	0	7	10	5	5	5	8	10
C7	9	10	8	5	10	10	5	5	4	3	15
C8	10	15	10	5	10	10	5	5	5	8	1
C9	9	15	10	5	10	10	5	5	2	2	10
C10	8	13	8	0	10	10	5	3.5	4	5	12
C11	8	15	7	0	10	10	2	5	5	3	12
C12	10	10	10	0	10	10	2.5	4	3	8	9
C13	10	8	10	0	10	8	3.5	4	4	9	10
C14	7	10	8	5	10	7	4	4	3	6	11
C15	10	10	2	0	10	10	5	4	3	6	15
C16	5	8	3	5	10	10	5	4	5	8	12
C17	8	15	1	5	10	8	4	4	2	5	10
C18	8	10	9	0	10	10	3.5	4.5	5	8	10
C19	10	15	10	0	7	9	4	4	5	8	1
C20	10	15	10	0	65	8	5	4	4	8	2
C21	7	11	10	0	5	9	4.5	4	4	8	3
C22	7	15	7	05	10	7	3.5	3.5	3	10	4
C23	8	15	8	0	7	8	2.5	3.5	5	10	4
C24	9	15	8	0	6	9	2.5	3.5	5	1	2
C25	2	15	2	5	5	10	5	4	2	1	12
C26	N	13	2	5	10	2	2	4	N	10	4
C27	8	15	8	0	6	6	2	5	5	4	10
C28	9	15	7	0	5	5	2.5	3	2	5	5
C29	8	13	8	0	7	9	2.5	4	5	10	1
C30	10	15	10	0	5	5	2.5	3	2	10	1
C31	9	15	8	0	5	7	2	3	4	5	10
C32	7	12	8	0	5	7	2.5	3.5	3	8	0.5
C33	7	12	7	5	5	6	3	3	3	5	1.5
C34	8	12	85	0	6	8	2.5	4	3	5	3
C35	8.5	15	5	0	5	5	2	3	3	0	4
C36	10	15	8	0	7	10	2	4	5	0	4
C37	8	10	5	0	10	8	5	4	5	10	4
C38	5	13	10	5	5	5	1.5	2	2	10	1
C39	8	15	8	0	5	3	1	2.5	3	2	10
C40	7	15	8	0	3	4	2	3	3	2	4
C41	7	12	5	0	5	3	1.5	3	3	1	3
C42	5	10	7	0	5	6	4	3	3	2	3
C43	5	10	5	0	7	5	2.5	3	3	5	3



C4	0.7	1	0.8	1	1	1	1	1	0.75	0.6	1
C5	0.8	0.8	1	1	1	1	1	1	1	0.9	0.66
C6	1	1	1	0	0.57	1	1	1	1	0.8	0.66
C7	0.9	0.67	0.8	1	1	1	1	1	0.83	0.3	1
C8	1	1	1	1	1	1	1	1	1	0.8	0.3
C9	0.9	1	1	1	1	1	1	1	0.33	0.2	0.66
C10	0.8	0.87	0.8	0	1	1	1	0.63	0.83	0.5	0.79
C11	0.8	1	0.7	0	1	0.25	1	1	1	0.3	0.79
C12	1	0.67	1	0	1	1	0.38	0.75	0.58	0.8	0.59
C13	1	0.53	1	0	1	0.8	0.63	0.75	0.75	0.9	0.66
C14	0.7	0.67	0.8	1	1	0.7	0.75	0.75	0.58	0.6	0.72
C15	1	0.67	0.2	0	1	1	1	0.75	0.58	0.6	1
C16	0.5	0.53	0.3	1	1	1	1	0.75	1	0.8	0.79
C17	0.8	1	0.1	1	1	0.8	0.75	0.75	0.17	0.8	0.66
C18	0.8	0.67	0.9	0	1	1	0.63	0.88	0.88	0.5	0.66
C19	1	1	1	0	0.57	0.9	0.75	0.75	1	0.8	0.03
C20	1	1	1	0	0.29	0.8	1	0.75	0.75	0.8	0.1
C21	0.7	0.73	1	0	1	0.9	0.88	0.75	0.75	0.8	0.17
C22	0.7	1	0.7	0	0.57	0.7	0.63	0.63	0.58	0.8	0.24
C23	0.8	1	0.8	0	0.43	0.8	0.38	0.63	0.92	1	0.24
C24	0.9	1	0.8	0	0.29	0.9	0.38	0.63	0.92	1	0.1
C25	0.2	1	0.2	1	1	1	1	0.75	0.17	0.1	0.79
C26	0.8	1	0.8	0	0.29	0.6	10.25	1	1	1	0.1
C27	0.9	1	0.7	0	0.57	0.5	0.38	0.5	0.25	0.4	0.66
C28	0.8	0.87	0.8	0	0.29	0.9	0.38	0.75	0.92	0.5	0.31
C29	1	1	1	0	0.29	0.5	0.38	0.5	0.33	1	0.03
C30	0.9	1	0.8	0	0.29	0.7	0.25	0.5	0.67	1	0.03
C31	0.7	0.8	0.8	0	0.29	0.7	0.38	0.63	0.58	0.5	0.66
C32	0.7	0.8	0.7	1	0.43	0.6	0.5	0.5	0.58	0.8	0
C33	0.8	1	0.8	0	0.29	0.8	0.38	0.75	0.58	0.5	0.07
C34	0.85	1	0.85	0	0.57	0.5	0.25	0.5	0.5	0.5	0.17
C35	1	0.67	0.5	0	1	1	0.25	0.75	1	0	0.24
C36	0.8	0.87	0.8	0	0.29	0.8	1	0.75	1	0	0.24
C37	0.5	1	0.5	1	0.29	0.5	0.13	0.25	0.25	1	0.24
C38	0.8	1	1	1	0	0.3	0	0.38	0.42	1	0.03
C39	0.7	0.67	0.8	0	0.29	0.4	0.25	0.5	0.58	0.2	0.66
C40	0.7	0.8	0.8	0	0.29	0.3	0.13	0.5	0.5	0.2	0.24
C41	0.5	0.67	0.5	0	0.57	0.6	0.75	0.5	0.5	0.1	0.17
C42	0.5	0.67	0.7	0	0.29	0.5	0.38	0.5	0.42	0.2	0.17
C43	0.5	0.67	0.5	0	0.29	0.5	0	0.25	0.33	0.5	0.1
C44	0.2	0.33	0.6	0	0.29	0.8	0.5	0.75	0.75	0.1	0.1
C45	0.2	0	0.2	0	0.14	0.6	1	0.25	0.33	0	0.52
C46	0.2	0	0.1	0	1	0	0.25	0.25	0	0	0.79
C47	0.2	0.13	0.2	0	0.71	0.5	0.38	0.25	0.42	0.1	0.03
C48	0	0.13	0	0	0.29	0.2	0	0	0.58	0.3	0.31



Now, with the methods developed, we embark on the analysis of the processed data. It is crucial to emphasize that this section is dedicated to ensuring precise data analysis and cluster ranking. The objective here is to establish a solid groundwork for constructing management perspectives and a more resilient customer management system, all rooted in the insights derived from this data.

**5. K-means, SILHOUETTE score, and clustering results and analysis**

In this section, to enhance confidence in the calculations and improve the analytical capabilities:

1. Clusters are investigated for a range of values, from  $K = 2$  to  $K = 10$ .
2. Due to the high sensitivity of this method to initial points, the clustering calculations have been repeated 50 times.

Moreover, in presenting the clustering results and silhouette scores, we have employed the average outcomes from multiple repetitions. The calculations were executed across 48 customer records, utilizing Euclidean distance across 11 measurement criteria and calculation scales. To illustrate, in the initial clustering calculation, the score results, as depicted in Table 6, are reported:

**Table 6.** The results of the first Clustering Repetition

Cluster	2	3	4	5	6	7	8	9	10
Silhouette scores	0.177	0.135	0.130	-0.031	0.068	-0.017	-0.111	-0.175	-0.061

The highest silhouette scores were achieved by clusters 2, 3, and 4, respectively. The progression of these scores can be observed in this calculation (Figure 1).



**Figure 1.** Results of first Clustering repetition

In the subsequent section, this calculation (similar to the process extracting Table 6) is repeated 49 more times, and the results are presented in Table 7. It will become apparent:

**Table 7.** Results of 50 Clustering Repetitions

Cluster	2	3	4	5	6	7	8	9	10
The mean of silhouette scores, 50 times repeated	0.177	0.135	0.108	0.060	0.044	0.000	0.046	0.068	0.091

The second cluster, with an average score of 0.177, exhibited the highest overall performance among all clusters. To facilitate a more comprehensive analysis, the results can be visualized in the corresponding graph

(Figure 2):

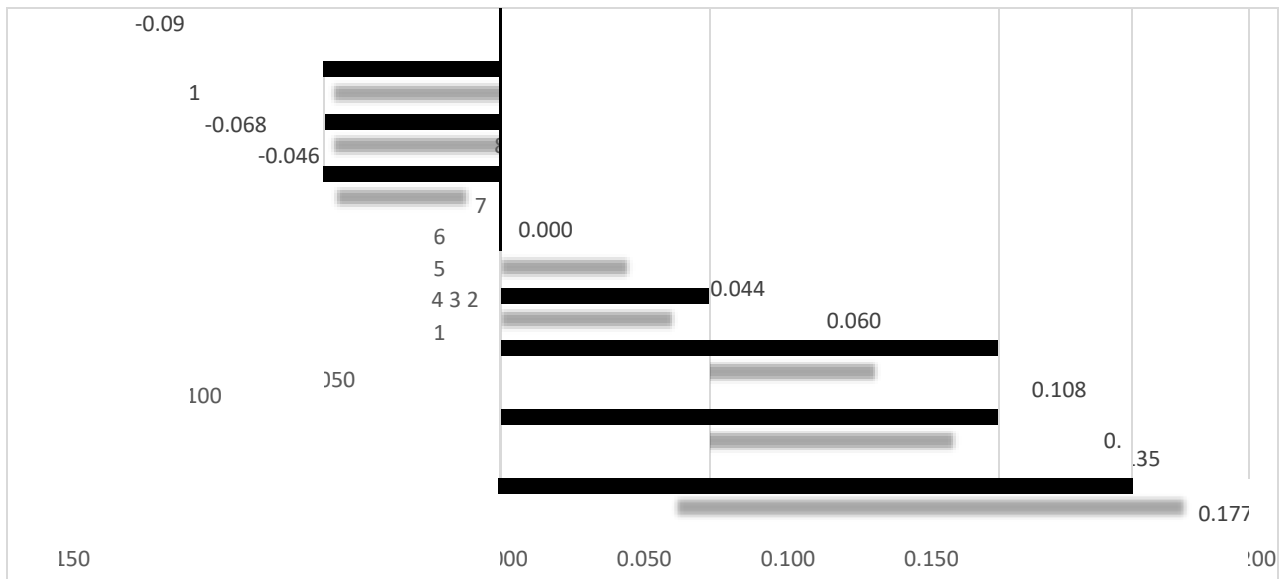


Figure 2. The results of 50 repetitions of clustering for each cluster (average silhouette score in 50 repetitions)

Interestingly, cluster 2 exhibited the highest mean silhouette score across 50 clustering repetitions. It is worth noting that, based on consultations and expert opinions from the desired manufacturer, segmenting customers into 2 clusters does not significantly create a competitive advantage in this industry. Therefore, it was recommended to consider the next optimal value, K = 3, for further investigation. In Table 8, we can observe the clustering results, including the clusters and the assigned customers:

Table 8. The results of Clustering based on clustering Categorization

Cluster	1	Cluster	2	Cluster	3
C1	C13	C8	C33	C45	C48
C2	C14	C19	C34		
C3	C15	C20	C35		
C4	C16	C21	C36		
C5	C17	C22	C37		
C6	C18	C23	C38		
C7	C25	C24	C40		
C9	C27	C26	C41		
C10	C31	C28	C42		
C11	C39	C29	C43		
C12	C46	C30	C44		
		C32	C47		

### 5.1. Fuzzy TOPSIS Analysis and Results

Now, with the clusters and assigned customers in hand, and utilizing the evaluation criteria outlined in Table 5, we can rank them. This involves assigning fuzzy points to each cluster in specific criteria to establish the cluster order. Below are the decision criteria along with their respective weights (Table 9).

**Table 9. Cluster evaluation criteria for fuzzy TOPSIS and corresponding fuzzy weights**

Cluster evaluation criteria	Average purchase share/percentage of total purchases	Average production power of the cluster	The arithmetic average Good Banking of the cluster
Criteria Weights	0.25	0.30	0.45
Criteria of Fuzzy Weights	(0.2, 0.3, 0.5)	(0.3, 0.4, 0.5)	(0.4, 0.5 0.7)

It is essential to highlight that, in assessing the clusters using fuzzy TOPSIS, the key criteria were extracted from the primary criteria in this study through insights gathered from experts within the relevant organization. Due to the equal importance attributed to these criteria, their weights are inherently identical, necessitating the assignment of matching linguistic expressions and fuzzy weights.

As indicated in Table 10, the focal point is the consideration of the average score within each cluster. Accordingly, for each numerical expression, a corresponding fuzzy expression is formulated and assigned to facilitate a comprehensive evaluation.

**Table 10. Non-Fuzzy Decision Matrix**

Non-fuzzy decision matrix	The average share of purchases in relation to total purchases	Average power in production and consumption	Arithmetic average of Good banking record
Cluster 1	0.44	0.79	0.83
Cluster 2	0.55	0.18	0.742
Cluster 3	0.17	0.43	0.15

As can be observed in Table 11, given that the responses were initially within a normal range before undergoing the averaging process, the interval [0, 1) encompasses our responses. Consequently, it becomes essential to partition this range into distinct categories and assign descriptive terms to them.

**Table 11. Allocation of Triangular Fuzzy numbers to Different parts of the answer range**

Triangular Fuzzy numbers	Descriptive statement	Limits of answer
(0.1- 0.2- 0.3)	Too weak	0-0.19
(0.3- 0.4- 0.5)	weak	0.2-0.39
(0.5- 0.6- 0.7)	normal	0.4-0.59
(0.7- 0.7- 0.8)	Better than normal	0.6-0.79
(0.8- 0.9- 0.9)	Good	0.8-0.89
(0.9- 0.9- 1)	Very good	0.9-0.99
(1- 1- 1)	Excellent	1

The development of a fuzzy decision matrix, followed by subsequent steps, plays a pivotal role in establishing cluster rankings. Table 12 intricately lays out the distances between the positive and negative ideals, complemented by the similarity index. This index effectively summarizes the outcomes of the ranking process achieved through fuzzy TOPSIS for each of the clusters.

**Table 12. Fuzzy TOPSIS results**

Alternative	D+	D-	CL	RANK
Cluster 1	0.165	1.815	0.917	1
Cluster 2	0.66	1.32	.667	2
Cluster 3	1.485	0.495	0.25	3

As highlighted in Table 12, Cluster 1 stands out with the highest similarity index, primarily attributed to its minimal distance from the positive ideal and maximal distance from the negative ideal. Subsequently, the clusters are ranked in descending order based on the similarity index. After determining the clusters, the customer relationship systems are designed in terms of four subjects: monetary credit allocation, discount amounts, the level of trust in selling products for each customer, and transportation cost. Note that in the past year of the investigated case study, these subjects have been considered experimentally without having a special system and framework. Hence, By analyzing the historical data of these subjects in the past years and taking into account parameters such as inflation, risk management, budgeting management, and forecasting the sales share of each cluster in the future, the company managers’ and stakeholders’ in the several meeting try to adjust the amounts of this subjects as shown in Table 13.

**Table 13. Defining customer relationship approaches regarding different clusters**

Subjects	The amount of monetary credit for each cluster	Discount amount of each cluster (\$ per kg)	The amount of trust selling product for each cluster (ton)	Transportation cost (\$ per Kg)
Cluster 1	150%*X*LME copper price	$Y < 0.06$	$100 < X < 200$	-----
Cluster 2	150%*X* LME copper price	$Y < 0.04$	$50 < X < 100$	0.001
Cluster 3	150%*X* LME copper price	$Y < 0.02$	$X < 50$	0.002

The allocation of the mentioned items to customers is instrumental in shaping effective communication within a defined system framework. This practice aims to streamline interactions with customers, offering tailored privileges based on their performance within their respective acquisition clusters. By implementing this approach, the engagement with customers becomes more personalized, allowing each customer to enjoy specific benefits and interactions with the organization following their performance in the designated clusters. For instance, a customer showcasing commendable performance in a specified criterion, such as maintaining a good account over a certain period, can be elevated to a higher cluster. Consequently, this customer is entitled to additional discount points and other predefined opportunities, as illustrated in Table 13. The proposed customer relationship system advocates for a win-win framework between customers and suppliers, fostering transparency in their working relationships. In this win-win framework, special points are allocated to customers exhibiting superior performance. Simultaneously, the supplier gains advantages from creditworthy customers who demonstrate reliability with fewer bounced checks, ensuring the realization of defined goals and strategies in the form of both cash and credit. This collaborative approach enables suppliers to have a clearer understanding of their customers' production plans and future endeavours.

**6. Conclusion**

In this research, a customer relationship system is proposed, employing a framework based on customer clustering through the combination of K-means and Fuzzy TOPSIS within a metal supply chain. Eleven distinct criteria were considered, and a real case study was conducted in the copper alloy production industry. The

identification of key criteria was initiated through interviews with main experts, leading to the preparation of data for the clustering technique after cleaning and processing.

The K-means method was leveraged to extract cluster ranking criteria, followed by the ranking of clusters using the Fuzzy TOPSIS method. Subsequently, after assigning customers to clusters and evaluating their performance against the criteria, strategies for each cluster were proposed and analysed. The cluster with the highest rank is designated a larger discount price range, among other factors outlined in Table 15. Similar strategies were formulated for clusters 2 and 3. This system contributes to the impartial treatment of customers, instilling discipline in the organization's activities through a rule-based approach. At a higher level, the proposal of a customer relationship system documents the organization's knowledge, which can be passed on to future generations of employees and managers.

Future developments may involve addressing uncertainty in criteria by clustering customers with fuzzy and probabilistic methods. Exploring alternative clustering algorithms, such as the DBSCAN clustering algorithm and Gaussian Mixture Model algorithm, presents another avenue for customer clustering. Additionally, integrating diverse and valuable methods for ranking through Multiple Attribute Decision-Making (MADM) methods like DEMATEL and PROMETHEUS is a promising direction. The model presented in this research holds applicability in other key industries, extending to contexts such as the food and petrochemical supply chain.

**Conflict of interest:** The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

## References

1. Afsar, A., Houshdar Mahjoub, R., & Minaie Bidgoli, B. (2014). Customer credit clustering for presenting appropriate facilities. *Management Researches in Iran*, 17(4), 1-24.
2. Andayani, U., Efendi, S., Siregar, N., & Syahputra, M. (2021). *Determination System for House Improvement Recipients In Serdang Bedagai By Using Clustering K-Means Method And Višekriterijumsko Kompromisno Rangiranje (Vikor)*. Paper presented at the Journal of Physics: Conference Series.
3. Ansari, A., & Riasi, A. J. I. J. o. B. (2016). Customer clustering using a combination of fuzzy c-means and genetic algorithms. *International Journal of Business Management*, 11(7), 59-66.
4. Bottou, L., & Bengio, Y. (1994). Convergence properties of the k-means algorithms. *Advances in neural information processing systems*, 7.
5. Gocer, F., & Sener, N. (2022). Spherical fuzzy extension of AHP-ARAS methods integrated with modified k-means clustering for logistics hub location problem. *Expert Systems*, 39(2), e12886.
6. Ikotun, A. M., Ezugwu, A. E., Abualigah, L., Abuhajja, B., & Heming, J. (2023). K-means clustering algorithms: A comprehensive review, variants analysis, and advances in the era of big data. *Information Sciences* 622, 178-210.
7. Jyothirmai, B., Rajendra, P., Kumari, D. A., & Aparna, N. High performance of Cluster-Based Strategy for reducing Delays in Wireless Sensor Networks. *Journal of the Maharaja Sayajirao University of Baroda*.
8. Kanungo, T., Mount, D. M., Netanyahu, N. S., Piatko, C. D., Silverman, R., & Wu, A. Y. (2002). An efficient k-means clustering algorithm: Analysis and implementation. *IEEE transactions on pattern analysis machine intelligence*, 24(7), 881-892.
9. Kaufman, L., & Rousseeuw, P. J. (2009). *Finding groups in data: an introduction to cluster analysis*: John Wiley & Sons.
10. Khadivar, A., & Mojiabian, F. (2022). Workshops clustering using a combination approach of data mining and MCDM. *Modern Researches in Decision Making*.
11. Kumar, S., Suhaib, M., & Asjad, M. (2021). Narrowing the barriers to Industry 4.0 practices through PCA-Fuzzy AHP-K means. *Journal of Advances in Management Research*, 18(2), 200-226.
12. Llamazares, B. (2019). Using interval weights in MADM problems. *Computers Industrial Engineering*, 136, 345-354.
13. Mahdiraji, H. A., Kazimieras Zavadskas, E., Kazemina, A., & Abbasi Kamardi, A. (2019). Marketing strategies evaluation based on big data analysis: a CLUSTERING-MCDM approach. *Economic research-Ekonomska istraživanja*, 32(1), 2882-2892.

14. Moradi Fard, M., Thonet, T., & Gaussier, E. (2020). Deep k-Means: Jointly clustering with k-Means and learning representations. *Pattern Recognition Letters*, 138, 185-192.
15. Moubayed, A., Injadat, M., Shami, A., & Lutfiyya, H. (2020). Student engagement level in an e-learning environment: Clustering using k-means. *American Journal of Distance Education*, 34(2), 137-156.
16. Namvar, M., Gholamian, M. R., & KhakAbi, S. (2010). *A two phase clustering method for intelligent customer segmentation*. Paper presented at the 2010 International conference on intelligent systems, modelling and simulation.
17. Ng, R. T., & Han, J. (1994). *E cient and E ective clustering methods for spatial data mining*. Paper presented at the Proceedings of VLDB.
18. Özari, Ç., & Can, E. N. (2023). Financial Performance Evaluating and Ranking Approach for Banks in Bist Sustainability Index Using Topsis and K-Means Clustering Method. *Academic Journal of Interdisciplinary Studies*.
19. Prahalad, C. K., & Ramaswamy, V. (2004). Co-creation experiences: The next practice in value creation. *Journal of interactive marketing*, 18(3), 5-14.
20. Raharja, M. A., & Surya, I. K. A. (2022). *Clustering Customer For Determine Market Strategy Using K-Means And TOPSIS: Case Study*. Paper presented at the Proceeding International Conference on Information Technology, Multimedia, Architecture, Design, and E-Business.
21. Rajagopal, D. S. (2011). Customer data clustering using data mining technique. *arXiv preprint arXiv:*
22. Razini, E., & Rasti, M. (2015). Assessing the competitive capacity of Iran's Copper Industry (Case Study of National Iranian Copper Company). *Quarterly Journal of Business Research*, 76, 81-51.
23. Sadeghi, M., Naghedi, R., Behzadian, K., Shamshirgaran, A., Tabrizi, M. R., & Maknoon, R. (2022). Customisation of green buildings assessment tools based on climatic zoning and experts judgement using K-means clustering and fuzzy AHP. *Building Environment*, 223, 109473.
24. Shahri, M. M., Jahromi, A. E., & Houshmand, M. (2021). Failure Mode and Effect Analysis using an integrated approach of clustering and MCDM under pythagorean fuzzy environment. *Journal of Loss Prevention in the Process Industries*, 72, 104591.
25. Syakur, M., Khotimah, B., Rochman, E., & Satoto, B. D. (2018). *Integration k-means clustering method and elbow method for identification of the best customer profile cluster*. Paper presented at the IOP conference series: materials science and engineering.
26. Valipour, M., Yousefi, S., Jahangoshai Rezaee, M., & Saberi, M. (2022). A clustering-based approach for prioritizing health, safety and environment risks integrating fuzzy C-means and hybrid decision-making methods. *Stochastic Environmental Research Risk Assessment*, 36(3), 919-938.
27. Wang, Y., Ma, X., Lao, Y., & Wang, Y. (2014). A fuzzy-based customer clustering approach with hierarchical structure for logistics network optimization. *Expert systems with applications*, 41(2), 521-534.
28. Wu, J., & Lin, Z. (2005). *Research on customer segmentation model by clustering*. Paper presented at the Proceedings of the 7th international conference on Electronic commerce.
29. Yankelovich, D., & Meer, D. (2006). Rediscovering market segmentation. *Harvard business review*, 84(2), 122.