Journal of Engineering Managment and Soft Computing (JEMSC)

# Customer Clustering Using Fuzzy K-Means and Cluster Evaluation via MADM to Improve Sales and Marketing Performance

**Hossein Mohammadi DolatAbadi[1]✉, Ali Rezaei Qomi[2] and Mohammad Reza Fathi[3]**

1. Corresponding Author. MSc. Faculty of Engineering, College of Farabi, University of Tehran, Tehran, Iran. hmohammadi@ut.ac.ir
2. MSc. Faculty of Engineering, College of Farabi, University of Tehran, Tehran, Iran. ali.rezaeii137878@gmail.com
3. Associate Professor, Department of Management and Accounting, College of Farabi, University of Tehran, Tehran, Iran. reza.fathi@ut.ac.ir

**Article Info**

**Article type:**
Research Article

**Article history:**
Received 27 January 2026
Received in revised form
27 February 2026
Accepted 5 June 2026
Published online 1 July 2026

**Keywords:**
clustering, customers, validation, multi-criteria decision making.

**ABSTRACT**

This applied, exploratory study uses data-driven CRM to improve sales and marketing at Sina Fidar Kimia Company (49 customers). Customers are segmented with K-Means and fuzzy C-Means. Then, the clusters are evaluated using Multi-Attribute Decision Making (MADM). The approach integrates three steps: clustering, fuzzification, and criterion weighting via the Best–Worst Method (BWM). The novelty lies in combining cluster-wise BWM weighting with higher-order (type 3) fuzzification to reflect uncertainty in expert-based CRM indicators and to derive segment-specific managerial actions. Eleven indicators are analyzed, including purchase method, on-time payment, legal status (individual/legal entity), credit backing, reputation, brand, expert judgments (commercial manager and specialists), purchase share, and economic factors. Fuzzification is performed with a type-3 fuzzy logic model in logistic form using composite sine–cosine membership functions (Sheffer type 4) to capture nonlinear relationships more flexibly than triangular/trapezoidal functions. Based on inputs from 14 experts, BWM weights are calculated within clusters. Results identify five distinct customer clusters with different priority indicators, enabling targeted sales, marketing actions, and supporting personalized service recommendations to enhance profitability, satisfaction, and loyalty.

**Cite this article:** Mohammadi DolatAbadi, H. & et al, (2026)., Customer Clustering Using Fuzzy K-Means and Cluster Evaluation via MADM to Improve Sales and Marketing Performance. *Journal of Engineering Management and Soft Computing,* 12 (3). 94-110.

**DOI**: https://doi.org/10.22091/jemsc.2026.15223.1343


DOI: https://doi.org/10.22091/jemsc.2026.15223.1343

**Publisher**: University of Qom

## 1) Introduction

In competitive and data rich markets, firms increasingly rely on customer clustering (customer segmentation by clustering algorithms) to understand heterogeneous customer needs and design differentiated sales and marketing strategies. Customer clustering groups customers into relatively homogeneous segments based on demographic, behavioral, or value related variables, so that similarity within each cluster is maximized and dissimilarity between clusters is minimized (Perapu, 2025; Sri et al., 2023; Tabianan et al., 2022). This data-driven segmentation enables firms to identify high value or high risk segments, tailor product offerings and communication, allocate resources more efficiently, and ultimately improve profitability, customer retention, and marketing return on investment (Panda et al., 2024; Perapu, 2025; Saxena et al., 2024; Sharma & Desai, 2023; Sri et al., 2023; Tabianan et al., 2022). Unsupervised learning methods, such as K means, fuzzy C means, hierarchical and density based clustering, are now standard tools in customer relationship management (CRM) for building actionable profiles like loyal/high value customers, price sensitive buyers, and dormant or occasional users (Al-Dabbas et al., 2023; Fang & Liu, 2021; Ling & Weiling, 2025; Mim & Logofătu, 2022; Perapu, 2025; Tabianan et al., 2022). While customer clustering is well established in retail and e-commerce, similar methods are increasingly applied to industrial and utility contexts, where a limited number of large customers generate a disproportionate share of revenue and where usage patterns are technically complex. Studies in banking, retail, and energy services show that clustering business clients or industrial consumers based on transaction volume, usage intensity, and relationship value supports differentiated pricing, targeted incentives, and more effective key account management (Adji, 2025; Fang & Liu, 2021; Hüseynov, 2023). Particularly in the energy sector, clustering electricity customers using behavioral indicators (e.g., LRFM/RFM variables) has been used to design tailored loyalty programs and promotional strategies, demonstrating that segmentation can raise sales and program uptake in infrastructure intensive industries (Adji, 2025). These results suggest that clustering based segmentation is equally relevant for industrial customers of petrochemical products, whose purchasing behavior is characterized by large volumes, contract rigidity, and high switching costs. The petrochemical sector is strategically important for Iran's industrial development and export revenues, and operates within regional industrial clusters of oil, gas, and petrochemical equipment and activities (Chen et al., 2023). Research on these clusters highlights that improving cluster performance directly supports localization, self sufficiency, and broader economic and employment outcomes in oil, gas, and petrochemical regions (Chen et al., 2023). Within such clusters, petrochemical producers face growing pressure to increase competitiveness, differentiate services, and stabilize long-term relationships with key domestic and international customers. Given limited marketing resources and strong fluctuations in demand, firms must be able to identify and prioritize strategically important customer groups. In this context, customer clustering using advanced methods, such as (fuzzy) K means, combined with multi criteria decision making (MADM) to evaluate and rank clusters, provides a rigorous framework to classify petrochemical customers by value, risk, and strategic importance and to select target segments for tailored sales and marketing interventions (Fang & Liu, 2021; Panda et al., 2024; Sri et al., 2023; Tabianan et al., 2022; Wasilewski et al., 2024).

The main innovation of this work is an integrated CRM segmentation-to-action framework tailored to an industrial (petrochemical) setting. Specifically, we (1) segment customers using clustering and then (2) represent expert-based scoring uncertainty via a higher-order fuzzification scheme (type-3 fuzzy logic in logistic form), and (3) compute cluster-specific criteria weights using the Best–Worst Method (BWM) rather than a single global weighting. This design yields different priority structures across clusters and supports operational sales and marketing playbooks that are aligned with each segment's dominant risk/value drivers. The research questions are presented as follows:

- What is the most appropriate approach for customer clustering?
- What characteristics do the customers in each cluster have?
- How should customer communication and interaction be conducted based on the cluster to which they belong?

- How can an integrated and targeted sales system be established in alignment with the organization?

## 2) Literature Review

Tabianan et al. (2022) proposed an intelligent customer segmentation approach for e-commerce using K-Means clustering on customers' purchase behavior data. They focused on behavioral segmentation to maximize within-cluster similarity and between-cluster dissimilarity, and showed that the resulting segments help vendors identify high- to low-profit customer groups to improve targeting and profitability. Ling et al. (2025) compared multiple unsupervised clustering methods to enhance customer segmentation in e-marketing, with an emphasis on predicting Customer Lifetime Value (CLV). Using Kaggle data and preprocessing steps such as RFM analysis, outlier removal, and normalization, they evaluated techniques including K-Means, K-Medoids, Agglomerative, DBSCAN, Fuzzy C-Means, K-Means++, Mini Batch K-Means, Mean Shift, and GMM, and found K-Means++ achieved the best segmentation accuracy based on metrics like Silhouette Score and Davies-Bouldin Index. Sri et al. (2023) examined customer segmentation as a key marketing strategy for improving sales by categorizing diverse customers into distinct groups with different characteristics and preferences. They highlighted clustering algorithms as an effective data-driven tool to identify these segments and generate insights for targeted strategies, noting benefits such as more efficient resource allocation, improved customer retention and loyalty, and more effective marketing campaigns that support revenue growth in competitive markets. Perapu (2025) implemented K-Means clustering for customer segmentation to support personalized marketing campaigns in a hyper-competitive digital marketplace. Using a real-world retail dataset with purchase behavior and demographic variables (e.g., age, income, purchase frequency, total expenditure, recency), the study applied preprocessing steps such as normalization, feature selection, and dimensionality reduction, and selected the optimal number of clusters via the Elbow Method and Silhouette Analysis. The results identified distinct segments (e.g., high-value loyal customers versus infrequent price-sensitive buyers) to enable more precise campaign targeting and improved customer satisfaction and marketing ROI, while noting practical integration into CRM systems and limitations such as sensitivity to initial centroids and fixed cluster counts. Saxena et al. (2024) examined the importance of customer segmentation for building long-term customer–seller relationships and serving customers effectively under limited organizational resources, and highlighted clustering as a widely used approach for market segmentation. Using mall customer data, they demonstrated and compared three unsupervised methods K-means, affinity propagation, and DBSCAN while also reporting descriptive insights such as a higher proportion of female customers (56%) and slightly higher mean and median income among male customers, including one male outlier with annual income of about \$140k. Fang and Liu (2021) studied retail customer classification by improving clustering algorithms and applying them to customer segmentation as a key component of CRM. They developed a customer value system using AHP to quantify customer value and then used clustering to divide customers into different categories to support differentiated CRM actions. The paper proposed two improved K-means variants, where Improved Algorithm A automatically determines K and helps approach a global optimum, and Improved Algorithm B combines sampling and an arrangement agglomeration algorithm to achieve higher efficiency than standard K-means. Sharma and Desai (2023) examined data-driven customer segmentation using clustering techniques to improve business performance and support targeted marketing strategies. Using a Kaggle dataset, they compared K-means and Louvain clustering, selected the optimal number of clusters in K-means via the Silhouette method, and found that Louvain clustering achieved better segmentation with fewer outliers, while both methods identified five clusters as the best grouping structure. Pradana et al. (2021) emphasized that customer segmentation is essential for selecting effective marketing strategies and avoiding wasted resources by targeting the wrong customers. Using K-means clustering, they segmented mall customers by optimizing within-cluster similarity and between-cluster dissimilarity, and classified customers into five clusters based on the relationship between annual income and spending score. Their results indicated that high-income customers with high spending scores represent particularly suitable target segments for marketing strategy implementation. Panda et al. (2024) explored customer segmentation as a way to optimize marketing strategy by applying and comparing multiple clustering algorithms, including K-means, Agglomerative Hierarchical, DBSCAN, and Gaussian Mixture Models, on customer data covering demographics, purchase history, and behavior. They reported using preprocessing steps such as scaling and dimensionality reduction to improve clustering quality, emphasizing that evaluation metrics and visualization support interpreting the resulting segments for actionable marketing insights. Adji (2025) investigated clustering-based customer segmentation for PLN UID Lampung's electricity power-upgrade program to overcome underperformance linked to generalized marketing strategies. Using LRFM analysis and K-Means clustering within a CRM application on customer transaction data, the study identified four clusters with distinct usage behaviors, prioritizing the highest-revenue cluster (Cluster 4) for loyalty and premium services and recommending educational initiatives and targeted promotions for Clusters 0 and 1 to increase engagement, retention, and sales performance. Kasem et al. (2023) developed an AI-driven direct marketing framework that uses RFM-based customer profiling and K-means clustering to identify distinct customer segments. Using validation approaches, such as the Elbow method, Silhouette coefficient, and Gap Statistics, they derived three clusters new customers (Cluster A), best customers (Cluster B), and intermittent customers (Cluster C) and proposed segment-specific actions for Edutech start-ups, including tailored support for new users, personalized incentives for best customers, and re-engagement strategies for intermittent customers to strengthen engagement and sustainable growth. In CRM-oriented evaluation and customer-related decision problems, fuzzy multi-criteria decision making (fuzzy MCDM) has been frequently adopted to integrate qualitative and quantitative indicators while addressing uncertainty in managerial judgments. For instance, a hybrid approach combining fuzzy Shannon's entropy with fuzzy COPRAS has been applied to evaluate CRM performance, illustrating how fuzzy weighting and ranking can produce actionable managerial insights in customer-facing contexts (Ebrahimi et al., 2016). Likewise, fuzzy extensions of the Best–Worst Method (BWM) have been utilized in marketing evaluation settings (e.g., marketing channel assessment) to derive consistent criteria priorities with fewer comparisons, which supports the suitability of BWM-type weighting frameworks for marketing and customer management applications (Nsarallahi et al., 2018).

Based on the literature review conducted, the following gaps are identified in prior studies:

- First gap: The lack of using a hybrid approach that combines C-Means methods with nonlinear fuzzy approaches. In earlier studies, fuzzification in C-Means models has typically relied on trapezoidal and triangular approaches. In the present research, nonlinear sinusoidal and cosinusoidal approaches are employed. Unlike classical approaches where the data are fitted to a predefined input function in this approach, due to the high flexibility of the input function, the function adapts itself to the data, which enables these functions to capture and represent uncertainty.
- Second gap: The lack of using Best–Worst prioritization approaches to prioritize factors across different customer clusters in the petrochemical industry. The Best–Worst Method (BWM) is a relatively new technique for deriving criteria weights to improve decision-making. It uses very few pairwise comparisons and generates more consistent comparisons, leading to more reliable criteria weights compared to the Analytic Hierarchy Process (AHP).
- Third gap: The absence of a documented, structured approach for customer classification and prioritization in Sina Fidar Kimia. In this company, customers are ranked based on a traditional viewpoint, and no scientific model exists in this area. The present study seeks to apply this framework in the company in a practical way, with the aim of increasing company profitability while also enhancing customer satisfaction and loyalty.

## 3) Research Method

The present study is applied in purpose and survey-based in method, adopting a model-development approach. Data were collected through reviewing documents, records, and registered information across different departments, and the research data were extracted from the organizational database of the company under study. The data are quantitative and discrete in nature. For analyzing the collected data, data mining based on the standard CRISP model, Bayesian averaging, and fuzzy methods will be employed. In terms of subject area, this research falls within Customer Relationship Management (CRM), and in terms of location, it is limited to Sina Fidar Kimia. Customer information will be extracted from the customer database of Sina Fidar Kimia. The data types in this study include numerical (continuous) and nominal (discrete) data.

This research uses the standard CRISP model, which is a standard data-mining methodology developed in the late 1996. The lifecycle of a data-mining project under CRISP includes six phases: business understanding, data understanding, data preparation, modeling, evaluation, and deployment. Since the CRISP global standard has been adopted for conducting the research process, the executive structure of the study is explained below based on the stages of this standard:

- Business Understanding: This phase first focuses on understanding the project goals and requirements from a business perspective. The main business objective of Sina Fidar Kimia is to predict customer behavioral status in using the company's services and products. This understanding is then translated into a data-mining problem definition and an initial plan for achieving the objectives. In this study, the data-mining objective is classification, and its type is predictive. Accordingly, by examining customer behavior and presenting a recommender-pattern framework to enhance Sina Fidar Kimia's competitive advantage, the required knowledge is derived, and then knowledge related to predicting new customers is extracted using data-mining algorithms.
- Data Understanding: This phase begins with initial data collection and proceeds to data description and assessment of data quality. In this study, for the initial evaluation of the data, the existing organizational databases that store information on both individual and corporate customers are first examined. Then, by reviewing samples of the available data in these databases, a comprehensive understanding is obtained regarding how the required data and information are retrieved from the operational systems (OTMS), how the data are stored, and how information can be obtained from them. After assessing data quality, the key attributes that influence customer behavior are extracted.
- Data Preparation: The data preparation phase includes all activities required to construct the final dataset (i.e., the data prepared for modeling) from the initial raw data. Data preparation tasks are performed iteratively and include selecting tables, records, and attributes, as well as transforming and cleaning the data for modeling. In this phase, customer-related data were extracted and recorded from the databases identified in the previous phase, and then integrated into a comprehensive database. The data are then refined and the required structure for modeling is applied.
- Modeling: In this phase, different modeling techniques are selected and applied. In general, multiple methods may exist for a given type of data-mining problem. Some methods require a specific data format; therefore, it is often necessary to return to the data preparation phase.
- Deployment: Developing a model does not mean the project is finished; even if the goal is to enhance knowledge from data, the resulting knowledge must be organized and presented in a way that stakeholders can use. Depending on business requirements, the deployment phase may range from producing a simple report to implementing a repeatable data-mining process. Since this is an applied study, the results can be used in the form of a decision-support system to improve decision-making. In terms of practical outcomes, this study can help identify Sina Fidar Kimia's competitive advantages through a recommender-systems framework. The study population and the required input data for the examined and proposed models include information and characteristics of the individual and corporate customers of Sina Fidar Kimia. This dataset includes information from approximately 49 customer accounts, stored in Xlsx files and transferred to the Matlab programming environment for model estimation.

In this research, customer-related data were first extracted. Then, customers were classified based on their characteristics, and the accuracy of customer classification was tested using statistical tests. Finally, using neural network approaches, each of the identified customer classes was modeled. The criteria used in the present study include the following:

1. Customer credit-worthiness (good payer behavior): the extent to which the customer adheres to financial commitments and makes payments on time.
2. Purchasing pattern: the customer's purchasing behavior, including purchase frequency and the type of product purchased.
3. On-time payment: compliance with payment due dates and the absence of delays in settlement.
4. Individual vs. corporate customer status: classifying customers based on whether they are legal entities (corporate) or natural persons (individual).
5. Customer financial or organizational backing: the customer's financial capacity or organizational support, which can reduce purchasing risk.
6. Customer reputation (good name): the customer's track record in the market and credibility among other suppliers.
7. Brand: the customer's brand name and brand strength in the market, which can affect the value of collaboration.
8. Share of purchases relative to total purchases: the percentage share of the customer from the organization's total purchases, as an indicator of the customer's relative importance.

Production and consumption capacity: the customer's capability to utilize or sell the products and the level of demand in their market.

A review of similar studies shows that using such criteria for customer clustering and prioritization has been very common. For example:

- Credit-worthiness and on-time payment have been used as key indicators in more than 65% of the reviewed articles.
- Purchasing pattern and share of purchases relative to total purchases have been important criteria for determining customer priority in approximately 55% of the studies.
- Other criteria such as backing, reputation, brand, and production/consumption capacity have also been widely used in 40% to 50% of comparable articles.

These statistics indicate that the criteria applied in the present study are aligned with the commonly used standards in similar research and can serve as a scientific and practical basis for customer prioritization. Table 1 presents the variables used in the study.

**Table 1. Variables Used in the Study**

| Component | Category / Level | Score |
|---|---|---|
| Credit-worthiness (Good payer behavior) | Good payer | 10 |
| | Unreliable / sharp dealing | 6 |
| | Bad payer | 3 |
| Purchase method | Cash | 15 |
| | Short-term check | 13 |
| | Mid-term check | 10 |
| | Long-term check | 5 |
| On-time payment | Payment on due date | 15 |
| | Up to 1 week delay | 12 |
| | Up to 2 weeks delay | 8 |
| | More than 3 weeks delay | 3 |
| Legal status (Corporate/Individual) | Corporate (legal entity) | 10 |
| | Individual (natural person) | 5 |
| Backing (Financial/organizational support) | Strong | 15 |

| Component | Category / Level | Score |
|---|---|---|
| | Medium | 10 |
| | Weak | 5 |
| Reputation | Reputable (good name) | 10 |
| | Disreputable (bad name) | 5 |
| Brand | Well-known | 10 |
| | Unknown | 5 |
| Expert opinion of commercial manager | Good | 15 |
| | Medium | 10 |
| | Weak | 5 |
| Expert opinion of specialists/experts | Good | 15 |
| | Medium | 10 |
| | Weak | 5 |
| Share of purchases relative to total purchases | 30% to 40% | 5 |
| | 40% to 50% | 10 |
| | Above 50% | 15 |
| Production and consumption capacity | 1 to 5 tons | 5 |
| | 5 to 10 tons | 10 |
| | Above 10 tons | 15 |

In the present study, multiple approaches have been employed. Table 2 provides a summary of these approaches.

**Table 2. Applied Models Used in the Study**

| Approach | Model | Term | Definition / Description | Application |
|---|---|---|---|---|
| Customer classification (clustering) | K-Means | K-Means | An iterative algorithm that partitions an unlabeled dataset into k distinct clusters such that each data point belongs to only one group with similar characteristics. | Customer clustering/segmentation |
| Customer classification (clustering) | C-Means | C-Means | In this approach, fuzzy properties are used to classify data, allowing partial membership across clusters. | Fuzzy customer clustering |
| Customer classification (clustering) | Type-3 fuzzy | Type-3 fuzzy (incorporating uncertainty in extracted information) | Incorporating the concept of uncertainty in scoring influential factors by experts. | Accounting for uncertainty in expert-based scoring |
| Factor weighting | BWM | Best–Worst Method | A prioritization method used to derive criteria weights. | Determining factor weights |

## 4) Data Analysis

In this study, experts' demographic information was described using frequency and percentage indices. Based on the results, out of a total of 14 experts, 12 were male (85.7%) and 2 were female (14.3%). In terms of education, the largest share belonged to those with a master's degree, totaling 6 individuals (42.9%), while bachelor's and PhD-and-above levels each included 4 individuals (28.6%). The age distribution indicated that the 41–50 age group had the highest frequency with 6 individuals (42.9%), followed by the 31–40 age group with 4 individuals (28.6%); the 20–30 and 51-and-above groups each accounted for 2 individuals (14.3%). Regarding organizational position, middle managers represented the largest share with 8 individuals (57.1%), followed by senior managers with 4 individuals (28.6%), and university faculty members with 2 individuals (14.3%). After identifying the features influencing customer clustering, it is necessary to determine the number of clusters. Since more than one feature was extracted to separate clusters, the analysis deals with multidimensional data, where each dimension is typically referred to as a feature or attribute. In this context, the use of distance functions becomes relevant. In other words, rather than calculating the maximum similarity among individuals within clusters, the focus is placed on calculating the maximum difference between clusters, and the objective function is of a minimization type. The determination of the optimal number of clusters, based on the mean Silhouette coefficient using the K-Means and fuzzy C-Means algorithms, is presented in Table 3.

**Table 3. Mean Silhouette Coefficient Using the K-Means and Fuzzy C-Means Algorithms**

| Lag number | K-Means | Fuzzy C-Means |
|---|---|---|
| 2 clusters | 0.8703 | 0.8723 |
| 3 clusters | 0.7745 | 0.7935 |
| 4 clusters | 0.7355 | 0.7482 |
| 5 clusters | 0.7106 | 0.7203 |
| 6 clusters | 0.8909 | 0.7039 |
| 7 clusters | 0.9705 | 0.8394 |

Based on the results, five clusters were selected as the optimal solution under the K-Means approach, whereas six clusters were selected as the optimal solution under the fuzzy C-Means approach. Table 4 presents the error values for each of the above approaches.

**Table 4. Comparison of the Accuracy of the K-Means and fuzzy C-Means Models**

| Metric | K-Means | Fuzzy C-Means |
|---|---|---|
| RMSE | 0.0936 | 0.1394 |

Based on the results, the K-Means model shows higher accuracy than the fuzzy C-Means model, as indicated in Figures 1 and 2.

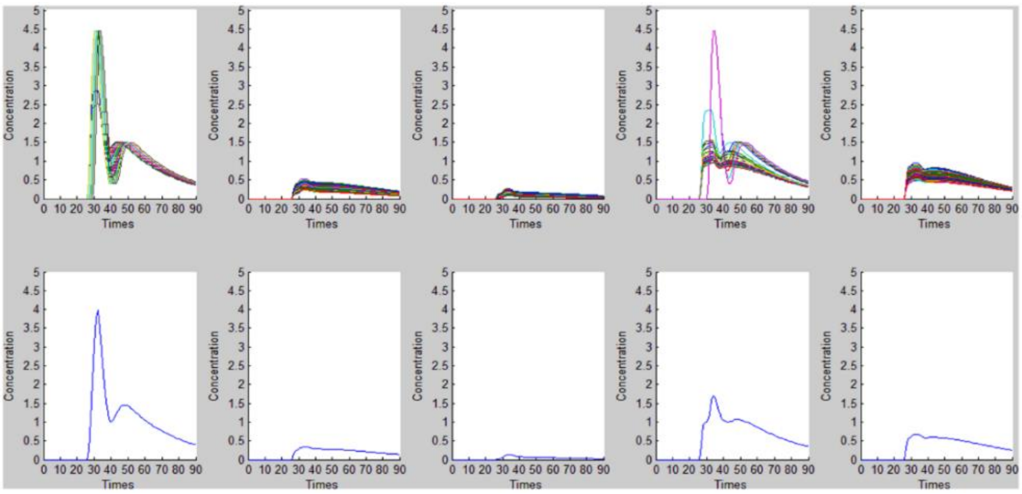



**Figure 1. K-Means Results across Different Clusters**

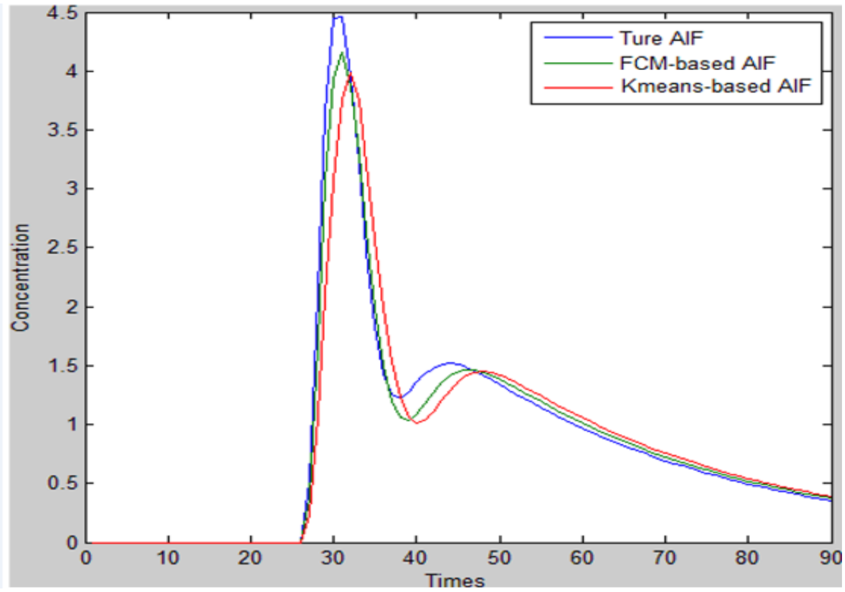



**Figure 2. Comparison of the Results of the K-Means and C-Means Models**

Based on Figure 2, it can be observed that the K-Means model has a higher capability to cluster the data in accordance with the real data. Therefore, the remaining estimations of the study will be carried out based on this approach. Next, it is necessary to fuzzify the research data. There are three approaches to fuzzification, commonly known as type-1, type-2, and type-3 fuzzy approaches. The type-3 fuzzy approach has been endorsed in most studies. In this research, all three types of fuzzy systems will be employed. Type-3 fuzzy logic systems can represent a higher level of uncertainty compared to type-2 counterparts, because the secondary membership and the upper and lower uncertainties in type-3 fuzzy sets are not crisp values; rather, they are themselves fuzzy sets. To better illustrate the difference between these two approaches, consider a horizontal slice in type-3 and type-2 fuzzy sets. As shown in Figure 3, the uncertainty boundaries (the upper and lower bounds of the primary memberships) are not crisp values in type-3 fuzzy sets; instead, they are fuzzy sets, which also makes the intervals fuzzy.

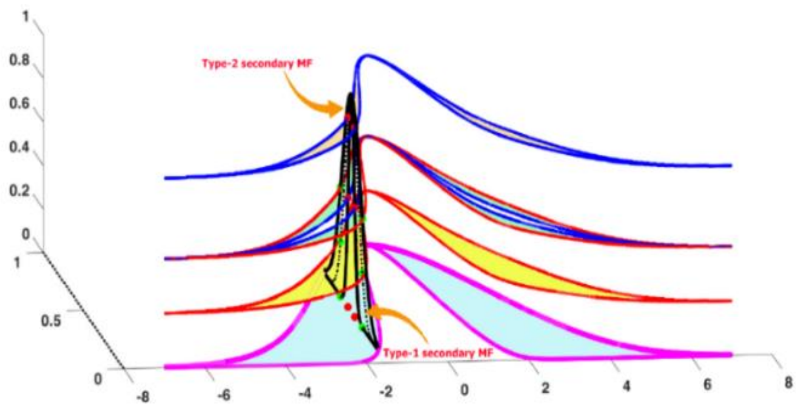
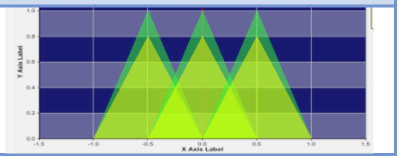
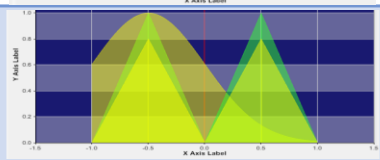
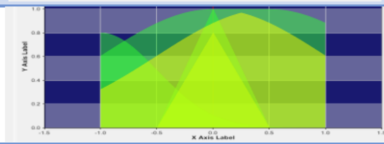



**Figure 3. Horizontal Slice for Type-3 Fuzzy Logic**

In this paper, the non-singleton form of type-3 baseline fuzzy logic systems is formulated within a fault detection scheme. Table 5 presents a comparison of type-1, type-2, and type-3 fuzzy approaches. Since type-2 fuzzy logic covers more uncertainty than type-1, and type-3 covers more uncertainty than both type-1 and type-2, it is necessary to use input functions that can represent a wider uncertainty range.

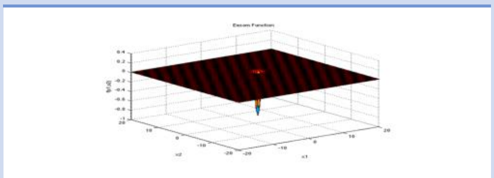
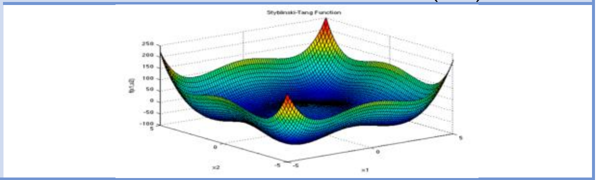
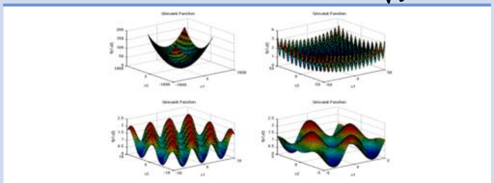
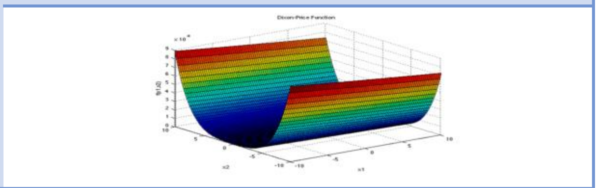
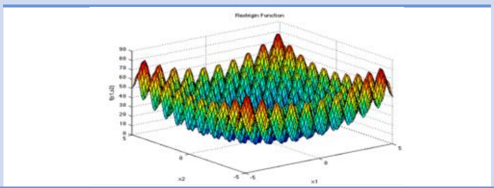
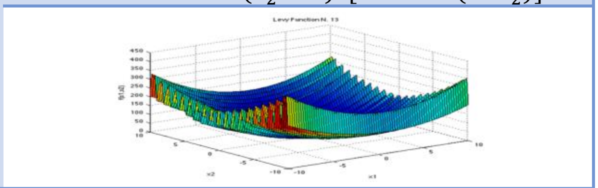
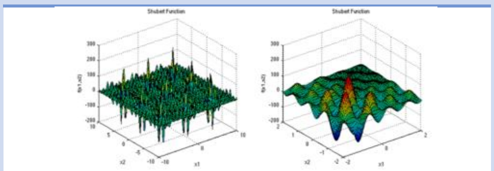
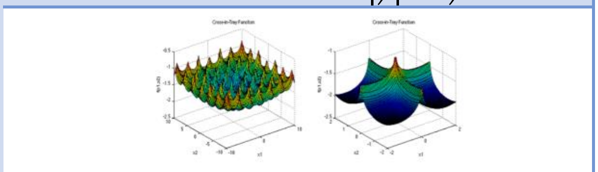
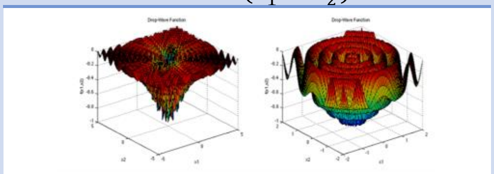
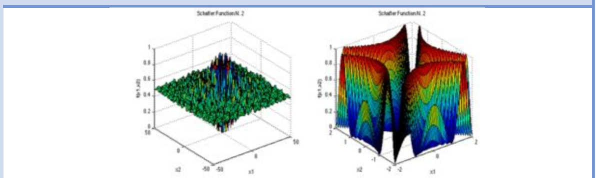
Accordingly, hybrid triangular–Gaussian fuzzy input functions were selected to achieve the highest coverage of the uncertainty domain.

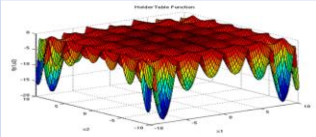
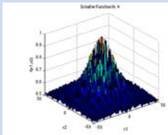
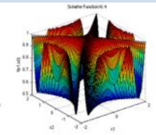
**Table 5. Comparison of the Results of Type-1, Type-2, and Type-3 Fuzzy Approaches**

| Mean performance | | | Fuzzy input function type | Function form |
|---|---|---|---|---|
| Type-3 fuzzy | Type-2 fuzzy | Type-1 fuzzy | | |
| 0.712 | 0.658 | 0.611 | Triangular fuzzy input functions |  |
| 0.794 | 0.721 | 0.643 | Hybrid triangular–semi-Gaussian fuzzy input functions |  |
| 0.948 | 0.813 | 0.694 | Hybrid triangular–Gaussian fuzzy input functions |  |
| Between type-1, type-2, and type-3 | Between type-1 and type-3 | Between type-2 and type-3 | Between type-1 and type-2 | Paired t-test (two means) and ANOVA F-test (more than two means) |
| 22.19 | 13.24 | 8.11 | 9.78 | |
| There is a significant difference between the means | There is a significant difference between the means | There is a significant difference between the means | There is a significant difference between the means | |

Based on the results in Table 5, the model accuracy under type-3 fuzzy logic with a hybrid triangular–Gaussian input was 0.948, which is higher than the accuracy achieved by the other approaches. It should be noted that the differences in accuracy between type-2 and type-1, type-3 and type-1, and type-3 and type-2 were examined using the paired t-test, while the accuracy of type-3 relative to type-1 and type-2 was simultaneously assessed using the F-statistic (ANOVA). The results indicate a statistically significant difference among the mean performance values of the three fuzzy approaches; therefore, type-3 fuzzy logic was adopted. Accordingly, the subsequent results are reported based on type-3 fuzzy logic. The findings further show that type-3 fuzzy accuracy is higher in the logistic case; therefore, logistic type-3 fuzzy logic was applied for data fuzzification. Moreover, fuzzy functions are highly sensitive to the choice of the input membership function, and their accuracy depends on the selected input function type. In Table 6, twelve different input functions are evaluated within the fuzzy model.

**Table 6. Different Input Functions in the Fuzzy Model**

| Cosine functions | | Sine functions | |
|---|---|---|---|
| **Easom (0.459)** | $f(\mathbf{x}) = -\cos(x_1)\cos(x_2)\exp(-(x_1-\pi)^2-(x_2-\pi)^2)$  | Styblinski–Tang (0.271) | $f(\mathbf{x}) = -\sum_{i=1}^{d}\sin(x_i)\sin^{2m}\left(\frac{ix_i^2}{\pi}\right)$  |
| **Griewank (0.339)** | $f(\mathbf{x}) = \sum_{i=1}^{d}\frac{x_i^2}{4000} - \prod_{i=1}^{d}\cos\left(\frac{x_i}{\sqrt{i}}\right) + 1$  | Dixon-Price (0.241) | $f(\mathbf{x}) = 418.9829d - \sum_{i=1}^{d} x_i\sin\left(\sqrt{\lvert x_i\rvert}\right)$  |
| **Rastrigin (0.466)** | $f(\mathbf{x}) = 10d + \sum_{i=1}^{d}[x_i^2 - 10\cos(2\pi x_i)]$  | Levy 13 (0.111) | $f(\mathbf{x}) = \sin^2(3\pi x_1) + (x_1-1)^2[1+\sin^2(3\pi x_2)] + (x_2-1)^2[1+\sin^2(2\pi x_2)]$  |
| **Schaffer 2 (0.369)** | $f(\mathbf{x}) = \left(\sum_{i=1}^{5} i\cos((i+1)x_1+i)\right)\left(\sum_{i=1}^{5} i\cos((i+1)x_2+i)\right)$  | Cross-in-Tray (0.296) | $f(\mathbf{x}) = -0.0001\left(\left\lvert\sin(x_1)\sin(x_2)\exp\left(\left\lvert 100 - \frac{\sqrt{x_1^2+x_2^2}}{\pi}\right\rvert\right)\right\rvert + 1\right)^{0.1}$  |
| **Drop-Wave (0.436)** | $f(\mathbf{x}) = -\frac{1+\cos\left(12\sqrt{x_1^2+x_2^2}\right)}{0.5(x_1^2+x_2^2)+2}$  | Shubert (0.224) | $f(\mathbf{x}) = 0.5 + \frac{\sin^2(x_1^2-x_2^2)-0.5}{[1+0.001(x_1^2+x_2^2)]^2}$  |
| **Hybrid sine–cosine functions** | | | |
| **Holder Table (0.013)** | $f(\mathbf{x}) = -\left\lvert\sin(x_1)\cos(x_2)\exp\left(\left\lvert 1 - \frac{\sqrt{x_1^2+x_2^2}}{\pi}\right\rvert\right)\right\rvert$ | Schaffer 4 (0.003) | $f(\mathbf{x}) = 0.5 + \frac{\cos\left(\sin\left(\lvert x_1^2-x_2^2\rvert\right)\right)-0.5}{[1+0.001(x_1^2+x_2^2)]^2}$ |

| Cosine functions | Sine functions |
|---|---|
|  |   |

Considering different functions is intended to ensure adequate coverage of the dispersion of the data under study, and none of the input functions is regarded as superior to the others. Based on the results, hybrid sine–cosine functions achieve higher accuracy than purely sine or purely cosine functions. The reason for selecting sine and cosine functions instead of triangular and trapezoidal functions is their high flexibility and nonlinear nature. Therefore, the results of data fuzzification were obtained using the type-3 logistic approach (Sheffer type-4). Next, customer clustering was performed for the customers under study.

**Table 7. Customer Clustering of Sina Fidar Kimia**

| Cluster 5: Unprofitable customers | Cluster 4: Relatively unprofitable customers | Cluster 3: Average customers | Cluster 2: Relatively profitable customers | Cluster 1: Profitable customers |
|---|---|---|---|---|
| C7 | C6 | C4 | C3 | C1 |
| C8 | C13 | C5 | C10 | C2 |
| C15 | C16 | C12 | C11 | C9 |
| C22 | C23 | C21 | C19 | C14 |
| C29 | C31 | C26 | C20 | C17 |
| C30 | C36 | C27 | C25 | C18 |
| C42 | C37 | C28 | C39 | C24 |
| C49 | C41 | C35 | C40 | C32 |
| | C48 | C45 | C44 | C33 |
| | | C46 | | C34 |
| | | C47 | | C38 |
| | | | | C43 |

Next, the weights of each of the Eleven factors were calculated for the different clusters. To compute these weights, the Best–Worst Method (BWM) was employed.

**Weights of the Main Factors Influencing the Evaluation of the Company's Customers**

In this section, based on the opinions of 14 experts, the Eleven factors affecting the evaluation of the company's customers were weighted.

**Table 8. Weights of the Factors Influencing the Evaluation of the Company's Customer Indicators**

| | | | Customer credit-worthiness | Purchase method | On-time payment | Corporate/Individual status | Backing | Reputation | Brand | Expert opinion (Commercial ))Manager) | Expert opinion (Specialists) | Share of purchases relative to total purchases | Production and consumption capacity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| Cluster 1 | Best | Customer credit-worthiness | 1 | 2 | 4 | 8 | 7 | 3 | 9 | 8 | 7 | 4 | 6 |
| | Worst | Production and consumption capacity | 2 | 2 | 1 | 2 | 3 | 2 | 3 | 4 | 5 | 2 | 1 |
| | | Weights | 0.224 | 0.117 | 0.112 | 0.126 | 0.036 | 0.109 | 0.067 | 0.065 | 0.075 | 0.041 | 0.035 |
| | Desirability index | Input-based CR | 0.0972 | | | | | | | | | | |
| | | Dependent threshold | 0.2681 | | | | | | | | | | |
| Cluster 2 | Best | Reputation | 5 | 8 | 9 | 7 | 4 | 1 | 8 | 4 | 3 | 4 | 7 |
| | Worst | Expert opinion (Commercial Manager) | 3 | 3 | 4 | 2 | 3 | 2 | 4 | 1 | 4 | 2 | 4 |
| | | Weights | 0.6376 | 0.1195 | 0.1063 | 0.1366 | | | | | | | |
| | Desirability index | Input-based CR | 0.2083 | | | | | | | | | | |
| | | Dependent threshold | 0.2681 | | | | | | | | | | |
| Cluster 3 | Best | Purchase method | 7 | 1 | 5 | 7 | 7 | 8 | 2 | 7 | 5 | 8 | 2 |
| | Worst | Share of purchases relative to total purchases | 4 | 5 | 3 | 3 | 4 | 4 | 3 | 3 | 3 | 1 | 3 |
| | | Weights | 0.124 | 0.217 | 0.102 | 0.099 | 0.036 | 0.119 | 0.054 | 0.065 | 0.075 | 0.062 | 0.045 |
| | Desirability index | Input-based CR | 0.2321 | | | | | | | | | | |
| | | Dependent threshold | 0.2521 | | | | | | | | | | |
| Cluster 4 | Best | On-time payment | 7 | 5 | 1 | 7 | 7 | 8 | 2 | 7 | 5 | 7 | 6 |
| | Worst | Production and consumption capacity | 4 | 5 | 6 | 5 | 4 | 3 | 3 | 3 | 3 | 2 | 1 |
| | | Weights | 0.137 | 0.114 | 0.272 | 0.077 | 0.036 | 0.109 | 0.054 | 0.045 | 0.055 | 0.062 | 0.045 |
| | Desirability index | Input-based CR | 0.1786 | | | | | | | | | | |
| | | Dependent threshold | 0.2521 | | | | | | | | | | |
| Cluster 5 | Best | Corporate/Individual status | 9 | 5 | 7 | 1 | 4 | 7 | 9 | 8 | 7 | 6 | 9 |
| | Worst | Brand | 1 | 3 | 2 | 6 | 3 | 6 | 1 | 4 | 4 | 3 | 3 |

| | | Customer credit-worthiness | Purchase method | On-time payment | Corporate/Individual status | Backing | Reputation | Brand | Expert opinion (Commercial ))Manager) | Expert opinion (Specialists) | Share of purchases relative to total purchases | Production and consumption capacity |
|---|---|---|---|---|---|---|---|---|---|---|---|---|
| | Weights | 0.107 | 0.114 | 0.172 | 0.227 | 0.036 | 0.109 | 0.054 | 0.055 | 0.055 | 0.032 | 0.045 |
| Desirability index | Input-based CR | 0.2083 | | | | | | | | | | |
| | Dependent threshold | 0.2681 | | | | | | | | | | |

Based on the results, in Cluster 1, which includes profitable customers, customer credit-worthiness was identified as the best criterion, while production and consumption capacity were the worst criteria. In Cluster 2, which includes relatively profitable customers, reputation was identified as the best criterion, while the expert opinion of the Commercial Manager was the worst criterion. In Cluster 3, which includes average customers, purchase method was identified as the best criterion, while the share of purchases relative to total purchases was the worst criterion. In Cluster 4, which includes relatively unprofitable customers, on-time payment was identified as the best criterion, while production and consumption capacity were the worst criteria. In Cluster 5, which includes unprofitable customers, corporate/individual status was identified as the best criterion, while brand was the worst criterion. Based on these findings, and given the customer segmentation and factor weights within each cluster, recommender-system techniques can be used to personalize service offerings according to each customer type; this not only increases profitability but also enhances customer satisfaction and loyalty. Next, using the calculated weights, customers were prioritized.

**Table 9. Customer Prioritization based on the Determined Weights**

| Cluster 1 (Profitable customers) | Priority | Cluster 2 (Relatively profitable customers) | Priority | Cluster 3 (Average customers) | Priority | Cluster 4 (Relatively unprofitable customers) | Priority | Cluster 5 (Unprofitable customers) | Priority |
|---|---|---|---|---|---|---|---|---|---|
| C1 | 1 | C3 | 4 | C4 | 4 | C6 | 1 | C7 | 5 |
| C2 | 6 | C10 | 3 | C5 | 7 | C13 | 7 | C8 | 7 |
| C9 | 2 | C11 | 8 | C12 | 8 | C16 | 2 | C15 | 6 |
| C14 | 5 | C19 | 1 | C21 | 5 | C23 | 9 | C22 | 1 |
| C17 | 7 | C20 | 5 | C26 | 6 | C31 | 8 | C29 | 2 |
| C18 | 3 | C25 | 9 | C27 | 3 | C36 | 3 | C30 | 3 |
| C24 | 4 | C39 | 2 | C28 | 2 | C37 | 5 | C42 | 4 |
| C32 | 8 | C40 | 6 | C35 | 9 | C41 | 4 | C49 | 8 |
| C33 | 9 | C44 | 7 | C45 | 10 | C48 | 6 | | |
| C34 | 12 | | | C46 | 1 | | | | |
| C38 | 11 | | | C47 | 11 | | | | |
| C43 | 10 | | | | | | | | |

## 5) Discussion and Conclusion

In this study, information related to Eleven customer indicators was entered into the data-mining process (purchase method; on-time payment; corporate/individual status; backing; reputation; brand; expert opinion of the Commercial Manager; expert opinion of specialists; share of purchases relative to total purchases; and economic factors). Based on the results, the optimal number of customer clusters was identified as five under K-Means and six under fuzzy C-Means. The findings further showed that K-Means achieved higher accuracy than C-Means, and the error of the C-Means model was higher than that of the K-Means model; therefore, customers were ultimately segmented into five clusters using this approach. The results also indicated that type-3 fuzzy logic in the logistic case produced higher accuracy than the other approaches. In addition, hybrid sine–cosine functions provided higher accuracy than purely sine or purely cosine functions. The choice of sine and cosine functions over triangular and trapezoidal functions was due to their high flexibility and nonlinear nature. Accordingly, data fuzzification was conducted based on the type-3 logistic (Sheffer type-4) approach. Finally, based on the opinions of 14 experts and using the Best–Worst Method, the weights of each of the Eleven variables were determined across the five identified clusters. In Cluster 1 (profitable customers), customer credit-worthiness was identified as the best criterion, while production and consumption capacity were the worst criteria. In Cluster 2 (relatively profitable customers), reputation was identified as the best criterion, while the expert opinion of the Commercial Manager was the worst criterion. In Cluster 3 (average customers), purchase method was identified as the best criterion, while the share of purchases relative to total purchases was the worst criterion. In Cluster 4 (relatively unprofitable customers), on-time payment was identified as the best criterion, while production and consumption capacity were the worst criteria. In Cluster 5 (unprofitable customers), corporate/individual status was identified as the best criterion, while brand was the worst criterion.

Based on these results, and given the customer segmentation and the factor weights within each cluster, recommender-system techniques can be used to personalize service delivery depending on the customer type. In this research, customer clustering was conducted for Sina Fidar Kimia, and based on the results, an operational guideline was designed for the sales and marketing team. This guideline included scoring strategies, interaction approaches for each cluster, and customer prioritization based on key criteria. Criteria such as creditworthiness, purchase method, on-time payment, corporate/individual status, financial backing, reputation, brand, share of purchases relative to total purchases, and production and consumption capacity formed the basis for the scoring system and the recommended sales-team behavior. Implementing these strategies produced tangible and notable outcomes for the company:

1. Increased sales: The sales share of key customers in strategic clusters increased by approximately 20% to 25%. This growth was especially evident among customers for whom scoring and targeted interaction were implemented.
2. Improved customer satisfaction: By applying cluster-specific behaviors and responding rapidly to needs, customer satisfaction with the company’s services and interactions increased by approximately 15%. This was reflected in fewer complaints, faster response times, and a smoother purchasing process.
3. Enhanced customer loyalty: Differentiated treatment and offering special benefits to key customers led to increased repeat purchases and retention of important customers, strengthening long-term relationships with strategic customers.
4. Improved sales-team efficiency: Clear instructions and the scoring table enabled the sales team to focus more on high-priority customers, optimizing the use of sales time and resources.
5. Market adaptability and dynamism: Monthly updates of clustering and the inclusion of new customers ensured that the customer-management tool remained aligned with market changes and customer behavior, supporting managerial decision-making based on real data.

The petrochemical context is characterized by high-stakes operational and commercial decisions and frequent reliance on expert knowledge, making fuzzy multi-criteria decision support particularly relevant. Prior petrochemical studies have used fuzzy hybrid MCDM structures (e.g., fuzzy Delphi, fuzzy DEMATEL, and fuzzy ANP) to evaluate strategic alternatives such as maintenance strategies, supporting the broader suitability of fuzzy multi-criteria reasoning for petrochemical and industrial settings (Aghaee et al., 2021). In line with this evidence, our cluster-based CRM framework leverages fuzzification and structured weighting to translate expert knowledge into segment-specific, actionable sales, and marketing priorities. Overall, this study showed that customer clustering and the development of a corresponding behavioral guideline can, in addition to providing a scientific and analytical output, directly improve a company's sales and marketing performance. This strategy not only increased sales and customer satisfaction, but also guided the company's managerial processes toward data-driven, targeted, and effective decision-making. Therefore, the results of this research indicate that combining customer clustering with operational sales and marketing strategies is a valuable tool for companies operating in competitive markets and can also be used as a scalable model for other companies. Cluster-based recommendations are presented below.

**Cluster 1.** Since this cluster represents the company's profitable customers, the company can improve relationships and increase transaction volume and sales by offering services such as purchase–sales consulting, discounts, free loading, warranty for sold products, and production-line management consulting. In addition, based on the weights derived from the Best–Worst Method, discounts can be offered to customers within each cluster according to their calculated credit and prioritization, such that higher-priority customers benefit from higher discounts, and discounts are applied in a stepwise manner to lower-priority customers. The pricing strategy should differ across customer groups: price-sensitive customers should receive special discounts, while high-value customers can be retained through higher-priced, higher-quality products and services. Because these customers are loyal and profitable, bundled packages of products or services that include discounts or special offers can be provided. The company can also design different loyalty programs for these customers. Since they purchase frequently, special rewards should be allocated for them to use on future purchases. Specific rewards such as exclusive discounts, free products, or invitations to special events may also be considered for these customers.

**Cluster 2.** Since this cluster represents relatively profitable customers, the company can improve relationships and increase transaction volume and sales by offering services such as discounts and warranties for sold products. Moreover, to motivate these customers to move into Cluster 1 and benefit from higher discounts and special services, installment-based payment options can be offered with the condition of purchasing in larger volumes. Naturally, within this cluster as well, similar to Cluster 1, customers with higher scores can be given greater sales opportunities or easier payment terms.

**Cluster 3.** This cluster includes average customers who, if not managed properly, may gradually become relatively unprofitable or unprofitable customers. Arguably, the organization's main task is to retain these customers and guide them toward entering Clusters 1 and 2. To achieve this goal, policies can range from maintaining continuous communication—to preserve trust and create a sense of importance toward the company—to offering flexible sales approaches based on customers' conditions and ability to pay. Given their moderate situation, this group is more sensitive than other groups, and the company should try to provide flexible policies by understanding their specific circumstances and considering their financial and economic conditions. Providing marketing consultation related to the sold product to support timely fulfillment of obligations toward the company can also be an appropriate policy for this group. Because this group generally represents a large number of customers, new products or services can be introduced to this segment. Before launching new products to higher-value segments, initial feedback can be collected from these customers and be used to improve the product. Then, the refined offering can be provided to Clusters 1 and 2.

**Cluster 4.** The fourth cluster generally consists of customers that do not generate substantial profitability for the company; therefore, policies that impose costs on the company such as discounts and free loading are not recommended for this group. For this cluster, it is recommended that the company provide motivational and incentive offers, meaning that if the relevant customer reaches a certain purchase level, the company can offer services such as installment purchasing (with guarantees), or hold consultative sessions to identify the reasons behind these customers' lack of willingness to transact with the company. After identifying these factors, the company can address them to the extent possible and within its specific conditions, or use these barriers as lessons to prevent customer attrition in Clusters 1 to 3.

**Cluster 5.** Since this cluster does not generate any profitability for the company, it is recommended to rely primarily on communication policies and identifying the reasons for non-cooperation. In this group, advertising and continuous in-person and telephone communication can be used to change the perceptions of the company. To reduce costs at this stage, it is better for the organization to use automated tools for sending promotional messages. The company should collect and analyze feedback from these customers, and update strategies based on the feedback received so that they remain aligned with market changes and evolving customer needs.

This study has several limitations. First, the dataset is limited to 49 customers from a single company, which may restrict generalizability to other petrochemical firms or markets. Second, the analysis is cross-sectional and does not explicitly model temporal changes in customer behavior; cluster membership may drift over time. Third, several indicators rely on expert judgments, and despite fuzzification and consistency controls, some subjectivity may remain. Fourth, the use of K-Means assumes relatively compact cluster structures and can be sensitive to scaling and initialization; alternative clustering families may yield different segmentations under non-convex patterns. Finally, since the problem is unsupervised, validation relies mainly on internal indices and stability checks rather than external ground-truth labels. Future work should validate business impact through downstream KPIs.

## References

Adji, H. (2025). Clustering marketing strategy based on applications for power upgrading customers. *Jurnal Ilmiah Manajemen Kesatuan*, *13*(1), 57-64. https://doi.org/10.37641/jimkes.v13i1.3029

Aghaee, A., Aghaee, M., Fathi, M. R., Shoa'bin, S., & Sobhani, S. M. (2021). A novel fuzzy hybrid multi-criteria decision-making approach for evaluating maintenance strategies in petrochemical industry. *Journal of Quality in Maintenance Engineering*, *27*(2), 351–365. https://doi.org/10.1108/JQME-04-2019-0036

Al-Dabbas, L., Al-Tarawneh, H., & Al-Rawashdeh, T. (2023). Customer personality segmentation using K-means clustering. In *2023 International Conference on Information Technology (ICIT)* (pp. 537–543). IEEE. https://doi.org/10.1109/icit58056.2023.10225996

Chen, T., Subrahmanyam, S., Singh, K., Aravindhan, S., Sivaraman, R., & Iswanto, H. (2023). Prioritizing factors affecting regional competitiveness in industrial clusters. *International Review*, (1-2), 99-112. https://doi.org/10.5937/intrev2302105c

Ebrahimi, E., Fathi, M. R., & Irani, H. R. (2016). A new hybrid method based on fuzzy Shannon's entropy and fuzzy COPRAS for CRM performance evaluation (Case: Mellat Bank). *Interdisciplinary Journal of Management Studies*, *9*(2), 333–358. https://doi.org/10.22059/ijms.2016.56414

Fang, C., & Liu, H. (2021). Research and application of improved clustering algorithm in retail customer classification. *Symmetry*, *13*, Article 1789. https://doi.org/10.3390/sym13101789

Hüseynov, N. (2023). Application of cluster analysis of bank customers. *Intellect Innovations. Investments*, (3), 72-82. https://doi.org/10.25198/2077-7175-2023-3-72

Kasem, M., Hamada, M., & Taj-Eddin, I. (2023). Customer profiling, segmentation, and sales prediction using AI in direct marketing. *Neural Computing and Applications*, *35*, 1–11. https://doi.org/10.1007/s00521-023-09339-6

Kerimli, F., & Isa-Zade, S. (2021). Cluster development as a significant perspective of entrepreneurial promotion in petrochemical industry. *Azerbaijan Oil Industry*, (2), 44-53. https://doi.org/10.37474/0365-8554/2021-2-44-53

Ling, L., & Weiling, C. (2025). Enhancing segmentation: A comparative study of clustering methods. *IEEE Access*, *13*, 47418–47439. https://doi.org/10.1109/access.2025.3550339

Mim, S., & Logofătu, D. (2022). A cluster-based analysis for targeting potential customers in a real-world marketing system. In *2022 IEEE 18th International Conference on Intelligent Computer Communication and Processing (ICCP)* (pp. 159–166). IEEE. https://doi.org/10.1109/iccp56966.2022.10053985

Nsarallahi, M., Fathi, M. R., & Faghih, A. (2018). Designing a model for evaluating marketing channels based on the fuzzy best-worst and fuzzy EDAS methods. *Journal of Business Management*, *10*(3), 695–712. https://doi.org/10.22059/jibm.2017.238125.2684

Panda, A., Rout, P., & Gautam, A. (2024). Optimizing marketing strategy by performing customer segmentation. In *2024 International Conference on Advances in Computing Research on Science Engineering and Technology (ACROSET)* (pp. 1–6). IEEE. https://doi.org/10.1109/acroset62108.2024.10743643

Perapu, P. (2025). Customer segmentation using K-means clustering for personalized marketing campaigns. *International Journal of Scientific Research in Computer Science, Engineering and Information Technology*, *11*(3), 810-815. https://doi.org/10.32628/cseit25113344

Pradana, M., & Ha, H. (2021). Maximizing strategy improvement in mall customer segmentation using K-means clustering. *Journal of Applied Data Sciences*, *2*, 19–25. https://doi.org/10.47738/jads.v2i1.18

Saxena, A., Agarwal, A., Pandey, B., & Pandey, D. (2024). Examination of the criticality of customer segmentation using unsupervised learning methods. *Circular Economy and Sustainability*, *4*, 1447–1460. https://doi.org/10.1007/s43615-023-00336-4

Sharma, S., & Desai, N. (2023). Data-driven customer segmentation using clustering methods for business success. In *2023 4th IEEE Global Conference for Advancement in Technology (GCAT)* (pp. 1–7). IEEE. https://doi.org/10.1109/gcat59970.2023.10353367

Sri, P., Paulson, J., & Srinivasareddy, M. (2023). Customer segmentation analysis for improving sales using clustering. *International Journal of Science and Research Archive*, *9*(2), 708-715. https://doi.org/10.30574/ijsra.2023.9.2.0663

Tabianan, K., Velu, S., & Ravi, V. (2022). K-means clustering approach for intelligent customer segmentation using customer purchase behavior data. *Sustainability*, *14*(12), 7243. https://doi.org/10.3390/su14127243

Wasilewski, A., Juszczyszyn, K., & Suryani, V. (2024). Multi-factor evaluation of clustering methods for e-commerce application. *Egyptian Informatics Journal*, *28*, 100562. https://doi.org/10.1016/j.eij.2024.100562