Volume 6, Issue 2, 062512 (1-12)

Fuzzy Optimization and Modeling Journal (FOMJ)

https://doi.org/10.57647/j.fomj.2025.0602.12


# A hybrid fuzzy framework for clustering and ranking listed companies via the K-means algorithm

Hamidreza Rezaei[1], Abdollah Hadi-Vencheh[2,*], Hosein Arman[3], Mehrdad Nikbakht[4]

[1] *Department of Management, Na.C., Islamic Azad University, Najafabad, Iran.*
[2] *Department of Mathematics, Isf.C., Islamic Azad University, Isfahan, Iran.*
[3] *Department of Management, Mo.C., Islamic Azad University, Mobarakeh, Isfahan, Iran.*
[4] *Department of Industrial Engineering, Na.C., Islamic Azad University, Najafabad, Iran.*

*Corresponding author: ahvencheh@iau.ac.ir

**Original Research**

Received:
2 May 2025
Revised:
3 June 2025
Accepted:
30 June 2025
Published online:
23 August 2025



**Abstract:**

The extant literature indicates that companies, including those listed on the Tehran Stock Exchange, are typically classified based on their respective industries. The findings of this study indicate that a reclassification of these companies may be warranted if alternative indicators are to be considered in the clustering process. To this end, the present study employs six indicators for the purpose of clustering the companies: price-to-return ratio (P/E), volume, liquidity rating, return on equity (ROE), $\beta$ coefficient, and return on capital. The values of these indices are extracted for each company, and then the companies are categorized into eight clusters using the K-means method. The significance of these groups varies from user to user, contingent upon the relative importance of the indicators from the user's perspective. To categorize these clusters, we initially employ the Fuzzy Analytical Hierarchy Process (FAHP) to ascertain the weights of the indexes. Subsequently, we utilize the TOPSIS method to prioritize the clusters. The results indicated that, according to the experts' assessments, the P/E ratio, beta coefficient, and liquidity rating, with weights of 0.365, 0.221, and 0.212, respectively, are the most significant indicators for the evaluation of the clusters. It is imperative to acknowledge that the decision matrix employed in this study is predicated on the assumption that the centroid of each cluster serves as its representative. The approach delineated in this study has the potential to serve as a foundational element in subsequent research endeavors focused on portfolio construction based on firm similarity.

**Keywords:** Clustering; K-means method; Hierarchical analysis process; TOPSIS method; Fuzzy sets

## 1. Introduction

In recent decades, the practice of active portfolio investment management has garnered significant acceptance among investors [1]. Investors often encounter difficulties in allocating capital effectively across various financial categories. Inadequate investment allocation has been demonstrated to have a deleterious effect on the return on investment. However, if the investment is distributed appropriately, portfolios can mitigate risk and enhance profitability. Consequently, the appropriate distribution of investments and the effective management of portfolios are imperative [2]. The selection of a stock portfolio constitutes a subject matter within the domain of finance. A portfolio is defined as the optimal combination of stocks or assets procured by an investor, with the objective of mitigating the overall risk of their investment while maintaining expectations regarding returns [3]. A portfolio is defined as a collection of assets owned by an institution or an individual. Portfolio management entails the strategic decision-making process regarding the acquisition, retention, and divestiture of both risky and risk-free assets over the course of an investment cycle, with these decisions being informed by the preferences of the investors [4]. A portfolio can comprise a combination of both similar and dissimilar assets. Therefore, a portfolio can be defined as a combination of different assets that may have different risk and return characteristics across components. Consequently, portfolio analysis is defined as the examination of the risk and return characteristics of assets within a portfolio, as well as the potential alterations that may arise from the interactions between these assets and their collective impact on other assets [5].

The evolution of financial markets has led to a diversification of investment opportunities, prompting individuals to allocate their savings across a range of financial categories in pursuit of enhanced profit potential. However, the financial market is a highly intricate system wherein investment opportunities and risks frequently coexist. For investors, it is imperative to implement effective risk management strategies while pursuing substantial returns. The financial markets are characterized by a perpetual state of flux, rendering it a formidable challenge to devise a portfolio investment strategy that can effectively account for this inherent volatility [6]. Portfolio analysis and selection is a term used to describe a set of techniques that decision-makers can utilize to make informed choices from a set of options. These techniques employ mathematical models that take into account constraints, preferences, and associated uncertainties.

The diversification of investment options has been a long-standing practice; however, the average option proposed by Markowitz has revolutionized the diversification of investment. This is a quantitative strategy, an approach that models the best assets [7]. A further contribution of Markowitz's work is the transformation of the financial decision-making process into an optimization problem. The model that was presented by the author of this study selects, among all possible portfolios, the one that presents the least risk for a given amount of return [8]. The Markowitz mean-variance model has historically served as the primary framework employed by numerous researchers in the selection of investment portfolios. However, it is important to note that this model is not without its limitations [3]. The most significant factor in constructing a portfolio is the projection of future stock market performance. The advent of sophisticated methodologies, novel forecasting capabilities, and portfolio selection methods has led to a proliferation of new approaches [9].

The prevailing discourse in the fields of financial calculus and stock selection places considerable emphasis on the evaluation of existing investments from the perspectives of risk and return rates. This enables investors to construct portfolios that align with their financial capacities and other strategic considerations. When presented with a range of investment opportunities, investors must determine the optimal number to select and the appropriate amount to allocate to each. This decision-making process can be quite complex under these circumstances. The final composition of a portfolio can result from two different processes. The first process is the result of random and unrelated choices. The second process is the result of careful planning by the investor [10]. In the domain of stock portfolio selection, decisions are predominantly guided by financial metrics. This reliance is predicated on the notion that a company's financial performance exerts a substantial influence on the determination of investment returns [11]. Research has demonstrated that factors such as liquidity, credit risks, and price volatility influence the appeal of corporate stocks [12]. Conversely, extant research has demonstrated that stock liquidity does not exert a significant influence on market efficiency or the prediction of stock returns on the Tehran Stock Exchange. This phenomenon can be attributed to the emergent nature of this market [13]. Multi-criteria decision-making (MADM) methods facilitate the comparison and ranking of portfolios by employing both quantitative and qualitative metrics. Consequently, portfolios are prioritized according to multiple established criteria, resulting in the selection of the optimal portfolio that meets logical constraints [14].

The optimization of stock portfolios is significantly facilitated by stock clustering in the capital market. A notable benefit of this approach is its capacity to facilitate the identification of stocks that are analogous or disparate based on the criteria delineated by the investor or analyst. This capability enables informed decision-making, as it provides a framework for evaluating and comparing potential investments with a high degree of clarity and precision. Another key benefit is the diversification achieved through stock clustering. A portfolio manager or investor who aims to manage investment risks typically seeks to diversify their stock portfolio to reduce unsystematic risk by holding a larger number of stocks [15]. The employment of a clustering technique facilitates the identification of efficient units, which are then distinguished from inefficient units. This process enables the development of a desired model, derived from the set of efficient units. The elimination of inefficient units reduces the number of units that contribute to the construction of the desired model, thereby enhancing the efficiency of the computational process related to the desired units [16]. The process of company classification, also referred to as industry classification, entails the systematic categorization of companies according to their commercial and industrial activities, as well as other pertinent factors. The objective of this classification process is to group companies with similar characteristics and differentiate them from others using various comparative parameters [17]. In the domain of financial research, the classification of companies is a prevalent methodology that entails the aggregation of analogous firms into discrete categories or clusters. This classification process has numerous practical applications for financial researchers, analysts, decision-makers, policymakers, and investors [18]. Homogeneity is a pivotal criterion for selecting the level of industry classification among available options and is frequently assessed using various methods. For instance, Chan et al. [19] advocate for the utilization of the co-movement of stock returns, while Bhojraj et al. [20] propose 12 key variables for classification. Phillips and Ormesby [21] developed a framework to understand the similarities and differences among firms by segmenting the market based on distinct business and financial characteristics.

Komeili and Birjandi et al. [22] proposed a methodology grounded in fundamental criteria to enhance information and minimize market uncertainties, thereby enabling the classification of companies within each industry. The objective of this study is to identify groups of companies that share similar market conditions and engage in comparable market activities. The stock exchange functions as an international financial marketplace for the exchange of securities. These markets function on a global scale, thereby playing a pivotal role in the global economy. Securities are a broad category encompassing a wide range of financial instruments, including stocks, bonds, and bank facility bonds, among others. These instruments are often traded on various stock exchanges, facilitating the purchase and sale of these assets. These markets wield considerable influence over the economic landscapes of nations and mirror the global economic situation. The Tehran Stock Exchange, also referred to as the Iran Stock Exchange, is a prominent financial market in the Middle East. The history of the Iranian stock exchange dates back to 1347, when it was founded with the objective of fostering a dynamic and active financial market within the country. Over the centuries, it has evolved into the preeminent hub for trading securities, shares, and a wide array of financial instruments in Iran. The Tehran Stock Exchange comprises multiple departments, including the Technology Development Department, the Iran Commodity Exchange, the Agricultural Commodity Exchange, and the Energy Exchange. Collectively, these entities offer investors a diverse array of investment opportunities, encompassing a wide range of securities and financial products. The market's adherence to international standards, as evidenced by its codified regulations, fosters a secure and lucrative investment environment for both domestic and foreign investors. The Tehran Stock Exchange, Iran's primary financial market, is involved in numerous industrial sectors. The Iran Stock Exchange is home to a diverse array of prominent industries, including steel, automotive, and petrochemical sectors. The presence of a wide array of industries and companies within the Tehran Stock Exchange highlights the dynamic nature of Iran's capital market.

The objective of this study is to categorize the companies listed on the Tehran Stock Exchange according to criteria defined by users and subsequently to rank these categories based on user preferences. This approach facilitates the establishment and classification of optimal clusters in accordance with user-defined indicators. The industry in which a company operates does not constitute a primary factor in this clustering; rather, it is determined by the users' perspective. It is important to note that other factors may hold more significance. For instance, certain users have posited that company size may be a more significant factor in clustering than the industry to which the company belongs. Consequently, this study first determines the user-selected indicators for grouping companies in the Tehran Stock Exchange and applies the K-means method to cluster them accordingly. The study's primary objective is to establish a hierarchical ranking of the clusters formed based on user input. To achieve this objective, the relative importance of the indicators is calculated using the Fuzzy Analytic Hierarchy Process (FAHP). The FAHP method is employed because experts typically express preferences for indicators in approximate terms, not exact values, and fuzzy sets effectively capture these estimations and the linguistic nuances used by experts. A decision matrix is subsequently formulated, with each cluster's center representing the entire group, and its indicators are entered into the matrix. Ultimately, the Technique for Order Preference by Similarity to the Ideal Solution (TOPSIS) is employed to prioritize the resulting clusters.

The contributions of this study are as follows:

- The objective of this study is to obtain the weights of clustering indicators using the FAHP method.
- The objective is to cluster the companies based on their similarities, as opposed to their industries.
- The center of each cluster is designated as its representative.
- The subsequent step involves the application of the TOPSIS method to assign a ranking to each cluster.

The remainder of this paper is organized as follows: Section 2 reviews the relevant background, section 3 details the research methodology, section 4 demonstrates the application of the proposed approach to a real-world case, and section 5 presents the discussion and concluding remarks.

## 2. Theoretical fundamentals and background research

### 2.1 Portfolio

Individuals perpetually endeavor to identify lucrative investment opportunities. Consequently, individuals perpetually seek out advantageous economic climates in which to invest. As a result, financial markets offer alluring prospects for the accumulation and generation of wealth. The acceptance of risk is a critical component of achieving profit. The management of these risks and uncertainties is contingent upon the possession of adequate knowledge and skills. Consequently, a pivotal challenge in financial management is the effective oversight of the investment process, which is frequently referred to as capital portfolio optimization or portfolio optimization [8]. Two fundamental criteria must be taken into account by all prospective investors. Firstly, it is imperative that the investment prioritize the maximization of returns. Secondly, the returns should be consistent, meaning that the investment risk should be minimized as much as possible [23]. Stock trading has emerged as one of the most prevalent investment strategies employed by investors to generate financial gains. A portfolio is defined as the group of stocks that an investor selects for analysis, monitoring, allocation, and reallocation. Investing in stocks invariably entails a certain degree of risk due to the inherently unpredictable nature of stock market price fluctuations. This indicates that, while the potential for high returns exists, there is also a possibility of loss or underperformance relative to expectations. Therefore, an investor's objective is to identify an investment strategy that aims to maximize profits while minimizing risk. Portfolio allocation is a strategy that considers time, investment objectives, and the level of risk [24]. The selection of a portfolio is not without its uncertainties, and researchers have proposed a variety of methodologies to address these concerns. For instance, Omidi et al. [25] employed the Mean-AVaR-skewness-kurtosis method to address the concerns of investors in uncertain fields. Similarly, Darabi et al. [26] utilized the Caputa-Clayton function to optimize portfolios and compared the outcomes of employing this method with those of the Markowitz model.

Markowitz's Modern Portfolio Theory remains a popular choice among investors due to its numerous practical advantages. The efficient frontier model is employed to assist investors in constructing a diversified portfolio and selecting a more efficient one.

Modern Portfolio Theory (MPT) serves as a foundational framework for research in quantitative portfolio allocation, facilitating the construction of asset portfolios that are aligned with investment objectives. These objectives may include either the maximization of returns within a specified risk level or the minimization of risk for a given expected return. A fundamental insight derived from Markowitz's contributions is that asset diversification strategies result in enhanced allocation outcomes. This concept utilizes statistical measures, such as variance and correlation, to demonstrate that the influence of investment on the overall portfolio outweighs its performance. For risk-averse investors, the risk contribution of each security in the portfolio is a critical factor [24]. The objective of selecting the optimal investment portfolio is to identify the ideal combination of stocks, assets, or securities to achieve the best outcomes based on specific criteria. The ability to make optimal investment decisions necessitates access to a more extensive array of data and a more profound comprehension of the factors that influence selection. In order to make more informed choices about stocks in the capital market, it is essential to understand the effects, and most importantly, the priority and significance of each influencing factor. The role of information in forecasting stock returns has prompted researchers to seek out variables and indicators that can explain stock performance [23]. A review of the extant literature on portfolio optimization has been conducted, and the results of the research have been presented in the form of a review article [27].

### 2.2 Clustering

Cluster analysis, also referred to as clustering, is the process of dividing a dataset or observations into smaller groups or subsets. Each subset is designated as a "cluster," signifying a group of data points that share commonalities while exhibiting distinctions from other clusters. The result of a cluster analysis, which consists of various clusters, is referred to as "clustering."

It is important to note that the implementation of different clustering techniques on the same dataset can result in the generation of distinct clusters. Due to the substantial scale and intricacy of data, manual execution of clustering is impractical for humans; therefore, the utilization of clustering algorithms becomes imperative. Clustering is a critical step in data analysis, as it facilitates the identification of previously unknown patterns or groups within the data [28]. The objective of cluster analysis is to identify the optimal arrangement of the data, ensuring that data points within the same cluster are homogeneous while being distinct from those in other clusters. Criterion functions are employed to assess the uniformity within clusters and the disparities between them [22]. The K-Means algorithm is classified as an incremental technique and is among the most frequently employed clustering methods. The methodology is initiated with the selection of K points as the initial centers of the clusters. Subsequently, each point is assigned to the nearest cluster center according to the distance measure. Subsequent to the establishment of the clusters, the centers of each cluster are recalculated and updated. The algorithmic process is repeated iteratively until the cluster centers exhibit no further change [29]. Clustering techniques manifest in a variety of forms; however, they are generally categorized into two types: agglomerative (bottom-up) and divisive (top-down) clustering. In the context of agglomerative or bottom-up clustering, the initial approach involves treating each item as its own cluster, thereby establishing the total number of clusters equal to the number of items in the dataset. The algorithm, consistent with the objectives of the researchers, combines clusters with items that are proximate to one another, thereby generating new, more extensive clusters. Conversely, in divisive or top-down clustering, all items initially form one large cluster, which is then split into smaller groups based on their similarity, as defined by the researcher.

### 2.3 The related works

This section reviews the studies conducted to form a portfolio based on clustering. In the domain of finance and capital markets, researchers have employed a range of clustering techniques. A subset of these techniques have been utilized exclusively for the purpose of streamlining the decision-making process, while others have been developed with the objective of reducing the time required for solution implementation in machine learning and meta-heuristic algorithms. For instance, in their study, Song et al. [30] employed a divisive clustering method to categorize stocks in the NASDAQ index based on liquidity, which served as the common characteristic for stocks within each cluster. The researchers initiated their study with the hypothesis that the market in its entirety is regarded as a single entity. Subsequently, the researchers proceeded to assess the liquidity of all constituent stocks within the aforementioned cluster. They then calculated the correlation between each pair of stocks. Subsequently, the stocks were clustered and divided into distinct groups according to their correlation levels. In the pursuit of ensuring the validity of the correlation differences among stocks within each cluster, certain researchers, including Liu et al. [31], have adopted an iterative process. In such scenarios, the researchers do not dictate the number of clusters formed. Conversely, some researchers determine the desired number of clusters for their analysis and impose this into their models based on their research requirements. Consequently, the clustering algorithm reaches a conclusion and displays the clusters once it attains the predetermined number, irrespective of the data relationships within each cluster. Kumari et al. [32] applied this approach to the task of clustering stocks. The researchers implemented the K-means method, in which K denotes the number of predetermined clusters. In the context of divisive or top-down clustering, all items are initially treated as a single cluster and subsequently divided into multiple clusters through a series of steps.

The fields of artificial intelligence, machine learning, and

data science have emerged as significant contributors to this transformation, offering numerous benefits and exerting a notable influence on the financial sector in recent years. A considerable number of researchers in the field of finance have utilized clustering techniques to categorize a wide range of goods, products, currencies, and associated items. One of the most prevalent applications of this methodology is the clustering of stock price time series. For instance, Durso et al. [33] analyzed time series by clustering according to time intervals and assessing their variances. In a similar vein, Tayali et al. [34] have proposed a methodology for the aggregation of time series based on the sequential arrangement of data points within the series. However, the application of clustering in finance extends beyond time series analysis; financial instruments are also grouped based on various characteristics and indicators, such as covariance and correlation coefficients. In their study, Leon et al. [35] initially computed the correlation coefficients for each pair of stocks and subsequently determined the final number of clusters using a divisive clustering approach.
The distinguishing characteristic of each cluster was the similarity in correlation coefficients among each pair of stocks within the clusters. Researchers generally seek to establish a criterion for measuring the distance between two stocks based on this feature. This allows for the assessment of the level of similarity among assets in each cluster.

## 3. Methodology

The primary objective of this research is to categorize the companies listed on the Tehran Stock Exchange according to the indicators valued by the study's participants. Additionally, the study seeks to evaluate and identify the top clusters based on their feedback. Given the integration of library research and field methods, including surveys, this research can be classified as analytical descriptive research in terms of its nature and approach. The research process is comprised of two primary phases. The initial phase is subdivided into four steps, while the subsequent phase is divided into three steps. The phases and steps of the research are illustrated in figure 1.
As illustrated in figure 1, the initial phase of this research focuses on grouping the companies that are listed. To achieve this objective, the use of multiple indicators is recommended. A broad array of these indicators exists, encompassing financial, profitability, and operational ratios, as well as purchase volume, return on assets, and numerous others. The selection of indicators is contingent upon the insights of the relevant experts or investors involved in the analysis and clustering of stock companies. In this study, a range of indicators was presented to participants, resulting in the selection of six key indicators for clustering the listed companies: price-to-earnings ratio (P/E), volume, liquidity rating, return on equity (ROE), $\beta$ coefficient, and return on capital. Subsequently, the values of these indicators were obtained for companies listed on the Tehran Stock Exchange. It is imperative to note that the purposive sampling method was employed, and only those companies for which the indicators could be measured during the study were included. The final sample comprised 332 companies. In this phase, the companies were classified into various clusters using the K-means method. The subsequent phase of this research entails the utilization of the K-means method for the purpose of ranking the formed clusters. To accomplish this objective, Multi-Attribute Decision Making (MADM) methods can be applied; this research employs a combination of Analytical Hierarchy Process (AHP) and Technique for Order Preference by Similarity to Ideal Solution (TOPSIS) methods. Specifically, the study first determines the

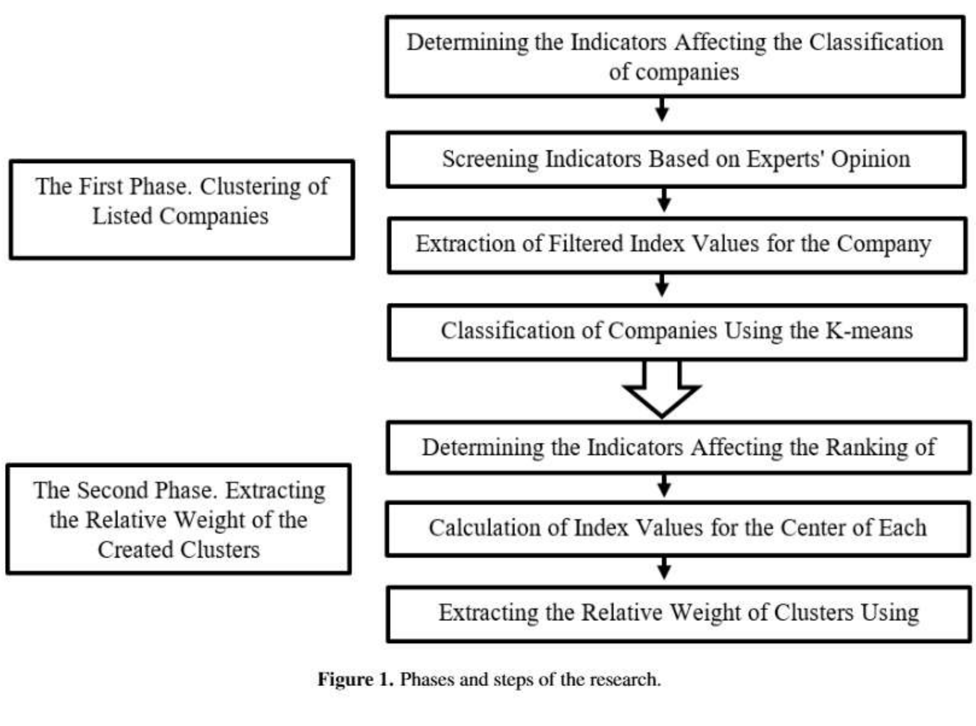


**Figure 1.** Phases and steps of the research.

weights of the indicators based on experts' opinions using the AHP method. Subsequently, the weights of the clusters are calculated using the TOPSIS method, and the clusters are then ranked accordingly. It is noteworthy that, given the tendency of experts to present their opinions in a relatively approximate manner, this research employs the fuzzy version of the AHP method to address this uncertainty and to quantitatively represent expert opinions.

A variety of models have been introduced to enhance the AHP method in a fuzzy context. The most well-known of these is the developmental analysis method. However, this method has significant shortcomings that can result in the incorrect assignment of weights to the alternatives. In some cases, this can lead to inaccurate rankings of the alternatives. Arman et al. examined the limitations of developmental analysis methods and other traditional FAHP approaches, subsequently proposing new methods for FAHP that avoid the shortcomings of classical FAHP techniques [36]. The present study employs one of these methods, known as the fuzzy geometric mean method (FGMM). In this approach, experts' preferences are initially gathered in the form of linguistic expressions to facilitate pairwise comparisons of factors. Subsequently, these linguistic preferences are converted into triangular fuzzy numbers with the assistance of Table 1.

To use the FGMM method, first, a pairwise comparison matrix completed with triangular fuzzy numbers is considered as follows.

So that

$$\tilde{B} = (\tilde{b}_{ij})_{n\times n} = \begin{bmatrix} (1,1,1) & (l_{12},m_{12},u_{12}) & \cdots & (l_{1n},m_{1n},u_{1n}) \\ (l_{21},m_{21},u_{21}) & (1,1,1) & \cdots & (l_{2n},m_{2n},u_{2n}) \\ \vdots & \vdots & \ddots & \vdots \\ (l_{n1},m_{n1},u_{n1}) & (l_{n2},m_{n2},u_{n2}) & \cdots & (1,1,1) \end{bmatrix}$$

$$\tilde{b}_{ij} = (l_{ij},m_{ij},u_{ij}) = \tilde{b}_{ij}^{-1} = (\frac{1}{u_{ij}},\frac{1}{m_{ij}},\frac{1}{l_{ij}}),\ \ i,j=1,\cdots,n \tag{1}$$

Then, the geometric mean of each row of the matrix is calculated as follows:

$$\tilde{M}_i = \Big( (\prod_{j=1}^{n} l_{ij})^{\frac{1}{n}}, (\prod_{j=1}^{n} m_{ij})^{\frac{1}{n}}, (\prod_{j=1}^{n} m'_{ij})^{\frac{1}{n}} , (\prod_{j=1}^{n} u_{ij})^{\frac{1}{n}} \Big),\ \ i=1,\cdots,n \tag{2}$$

where $\tilde{M}_i$ is the geometric mean of triangular fuzzy preferences in row $i$ and $n$ dimensions of the matrix. In the following, it is normalized as follows:

$$\tilde{S}_i = \Bigg( \frac{\big(\prod_{j=1}^{n} l_{ij}\big)^{\frac{1}{n}}}{\sum_{k=1}^{n}\big(\prod_{j=1}^{n} u_{ij}\big)^{\frac{1}{n}}}, \frac{\big(\prod_{j=1}^{n} m_{ij}\big)^{\frac{1}{n}}}{\sum_{k=1}^{n}\big(\prod_{j=1}^{n} m_{ij}\big)^{\frac{1}{n}}} , \frac{\big(\prod_{j=1}^{n} u_{ij}\big)^{\frac{1}{n}}}{\sum_{k=1}^{n}\big(\prod_{j=1}^{n} l_{ij}\big)^{\frac{1}{n}}} \Bigg),\ \ i=1,\cdots,n \tag{3}$$

where $\tilde{S}_i$ is the fuzzy relative weight of the element corresponds to row $i$, and it can be considered approximately as a triangular fuzzy number. For dephasing in this research, the center of gravity method is used, whose formula for the triangular number is as follows:

$$S_i = \frac{l_i + m_i + u_i}{3} \tag{4}$$

where $S_i$ represents the diffused weight of the $i^{\text{th}}$ element. The sum of the obtained $S_i$ is not necessarily equal to 1. Therefore, they are normalized using the following formula so that their sum equals 1:

$$W_i = \frac{S_i}{\sum_{t=1}^{n} S_t},\ \ i=1,\cdots,n \tag{5}$$

where $W$ represents the exact weight of the $i^{\text{th}}$ element.

## 4. Evaluation and results

The subsequent analysis is structured in two phases. Initially, the companies are clustered, and subsequently, the clusters are ranked. A thorough elucidation of these two phases is provided herein. Additionally, a subsection was incorporated at the conclusion of the results section to evaluate the clustering quality and sensitivity analysis.

**Table 1.** Linguistic preferences and their equivalent triangular fuzzy numbers.

| Language preferences | Row-to-column priority | Column-to-row priority |
|---|---|---|
| Equal importance | (1, 1, 1) | (1, 1, 1) |
| Equal to relatively more important | (1, 2, 3) | (0.33, 0.5, 1) |
| Relatively more important | (1, 3, 5) | (0.2, 0.33, 1) |
| From moderately important to high importance | (3, 4, 5) | (0.2, 0.25, 0.33) |
| High importance | (3, 5, 7) | (0.14, 0.2, 0.33) |
| High importance to very high importance | (5, 6, 7) | (0.14, 0.17, 0.2) |
| Very important | (5, 7, 9) | (0.11, 0.14, 0.2) |
| Very high to absolutely important | (7, 8, 9) | (0.12, 0.13, 0.14) |
| Absolutely important | (7, 9, 9) | (0.11, 0.11, 0.14) |

### 4.1 The first phase clustering

In the preliminary phase, the research findings indicate the aggregation of listed companies into disparate clusters through the implementation of the K-means method. To this end, we initially ascertained the salient indicators for clustering. A plethora of indicators are utilized for the purpose of clustering, including financial ratios and performance ratios, among others. In this study, a set of indicators was presented to experts in the field and asked to identify the most significant indicators for the purpose of clustering. Consequently, a selection of six indicators was made based on the opinions of experts in the field: price-to-earnings (P/E) ratio, trading volume, liquidity rating, return on equity (ROE), beta coefficient ($\beta$), and return on capital. It is important to note that these indicators utilize disparate scales. For instance, the P/E ratio is a measure of valuation, whereas trading volume is based on a monetary unit. Consequently, the values of these indicators are not homogeneous. This may have implications for the subsequent clustering results. To circumvent this misapplication, the values corresponding to each indicator were normalized. The number of clusters may vary depending on the expert preferences; in this study, eight clusters were selected. The K-means clustering of these companies into eight clusters was performed using Minitab software, and the results are provided in Table 2.
As Table 2 shows, the number of companies categorized in clusters 1 to 8 is 1, 85, 1, 75, 6, 4, 103, and 57, respectively.

### 4.2 The second phase ranking the clusters

During the initial phase, the listed companies were grouped into 8 distinct clusters. This phase addresses the question of which cluster holds higher priority for users. To resolve this, first, the weights and significance of the 6 indicators are determined based on users' preferences using the FAHP method, followed by ranking the clusters via the TOPSIS method. The results outcomes of applying each technique are presented independently.
Obtaining the Indicators' Weights Using the FAHP Method. Here, the FGMM method is used to calculate the weight of the 6 indicators valued by the users. To achieve this, a user panel was assembled, where participants provided input on the pairwise comparison of indicators, achieving a group consensus. These qualitative comments were transformed into triangular fuzzy numbers, as outlined in Table 1, with the results displayed in Table 3.

**Table 2.** Eight clustering of companies based on the considered indicators.

| Cluster | Abbreviated names of companies |
|---|---|
| 1 | Shasta |
| 2 | Vatjarat, Pasa, Ghapino, Lakhzar, Kokhak, Lepars, Ghashsafa, Taira, Teksha, Druze, Zakotsar, Sakhzar, Kahmada, Kegaz, Shiran, Khafner, Kavir, Qpira, Klond, Ghazer, Ghaliber, Shevinde, Ghamhera, Qalrost, Webshahr, Feros, Veskab, Amin, Vandaf, Sasharq, Vakret, Beswich, Velmelt, Payt, Volkar, Qashhad, Dakimi, Bekab, Vibime, Chekapa, Decina, Vamdir, Kamengans, Ghabnoush, Pekarman, Lotus, Kanwar, Fakhas, Khosaz, Wasakht, Pardis, Kimia, Vamohan, Vetosem, Kapshir, Daabid, Sep, Ghadam, Shepna, Vepasar, Ghagol, Bemotu, Volghadar, Fazer, Kechad, Hamor, Khashrek, Sesafha, Saman, Zepars, Kegel, Chedan, Nimrino, Tekno, Khazar, Qamro, Kama, Vaafari, Fespa, Fasmin, Shiraz, Vaazer, Hatuka, Tanvin, and Ghadir |
| 3 | Khasapa |
| 4 | Labsa, Shepaksa, Khazamiya, Zob, Venvin, Moadan, Fasrab, Chafst, Khamhar, Kala, Famrad, Khakar, Floleh, Ghobshahr, Chekarn, Karazi, Shesina, Wekar, Shafars, Veniro, Energy, Gashkar, Vaneft, Asia, Thamskan, Ghasalem, Mellat, Kazer, Fajam, Thashehed, Ranfur, Omid, Dekotsar, Kamase, Damin, Vatosa, Dejabar, Darazak, Thanosa, Dezharavi, Katbas, Shalab, Kharing, Sasharq, Fasazan, Kassapa, Desbehan, Vebahman, Khamhor, Sekard, Laserma, Sabik, Silam, Sedor, Qazvin, Sahrmoz, Ghashad, Bekam, Shenft, Khreikht, Ghapak, Afog, Kafra, Pakshu, Opal, Faravar, Ghasabat, Khabahman, System, Dalqma, Lebutan, Khalent, Hafari, Rakish, Ghashan |
| 5 | Khazin, Abada, Plask, Abad, Kodma, Ghagerji |
| 6 | Automobile, Webmelet, Steel, Femli |
| 7 | Khagstar, Shetran, Shabandar, Vebasadr, Sabhsaz, Fakhuz, Farak, Petrol, Kimiatech, Bors, Hiweb, Up, Khatoqa, Vepost, Vakhrazm, Vesapa, Dalbar, Pars, Velsapa, Shabriz, Fars, Akhabar, Fayra, Vepars, Ghazer., Fama, Thamid, Khamharka, Asiatek, Shefan, Vakhavar, Rampana, Khan, Vaghdir, Vesina, Hatayd, Petair, Betrans, Ma, Ghannoosh, Tapiko, Hafars, Tfars, Qoran, Shaspa, Valsanam, Alborz, Kaveh, Atkam, Parsian, Midko, Shepars, Khansir, Tepampi, Kasram, Tamlat, Madaran, Vespe, Retap, Ghashazer, Parsan, Pedraksh, Vasanat, Khacharkhash, Ghakuresh, Khmutor, Shekharban, Qanisha, Ghadasht, Tasiko, Khatrak, Wamid, Dedam, Vetos, Beniro, Sepid, Wati, Valber, Vobo Ali, Ghachin, Vetoshe, Katram, Fanwal, Chafiber, Folage, Vetoka, Verna, Sita, Takemba, Veniki, Sidko, Kahafz, Sepaha, Karoi, Kepars, Simorgh, Kasadi, Beshab, Fanord, Kafpars, Kabafaq, Shamla, Qahkamat |
| 8 | Dana, Shepdis, Fasbezvar, Shagdir, Boali, Sharak, Jam, Ardestan, Nouri, Deler, Sefers, Defara, Vepetro, Tipico, Sarbil, Setran, Senir, Sarud, Sosofi, Mobeen, Shobhorn, Dashimi, Kermasha, Bafjer, Kesaweh, Shekler, Zamgsa, Detmad, Shefa, Segreb, Dabur, Daro, Sabjno, Sebhan, Sarum, Sefano, Jam Pilen, Depars, Sarab, Sekhouz, Haptero, Pashand, Shakharek, Samazen, Defra, Khorasan, Sekrama, Sedasht, Khederha, Dasoeh., Sehgamt, Sefar, Sakhash, Saqayin, Fpenta, Sharaz, and Vapkh |

**Table 3.** Table of pairwise comparisons of indicators.

| Indicators | P/E Ratio | Volume | Liquidity rating | ROE | Coefficient $\beta$ | Return on capital |
|---|---|---|---|---|---|---|
| P/E ratio | (1, 1, 1) | (7, 8, 9) | (1, 2, 3) | (3, 5, 7) | (1, 2, 3) | (3, 5, 7) |
| Volume | | (1, 1, 1) | (0.14, 0.17, 0.2) | (0.2, 0.33, 1) | (0.14, 0.17, 0.2) | (0.33, 0.5, 1) |
| Liquidity rating | | | (1, 1, 1) | (1, 3, 5) | (1, 1, 1) | (1, 3, 5) |
| ROE | | | | (1, 1, 1) | (0.2, 0.33, 1) | (1, 1, 1) |
| Coefficient $\beta$ | | | | | (1, 1, 1) | (3, 4, 5) |
| Return on capital | | | | | | (1, 1, 1) |

To ascertain the weights of the indicators, the geometric mean of each row's values is initially calculated according to formula (2). These values are then normalized using formula (3) and subsequently diffused using formula (4). Finally, the diffused values are normalized via formula (5), yielding the final weights of the indicators. The results of these calculations are presented in Table 4.

According to the findings presented in Table 4, experts have identified the price-to-earnings ratio, beta coefficient, and liquidity rating as the most critical indicators for cluster ranking, with weights of 0.365, 0.221, and 0.212, respectively.

The application of the TOPSIS method to the evaluation of cluster rankings is a subject of ongoing research. This section provides a comprehensive overview of the ranking of clusters generated through the TOPSIS method. To achieve this objective, the decision matrix is initially constructed. In this study, the center of each cluster is defined as the cluster itself, and the values of six indicators are calculated for each center to establish the decision matrix. To elaborate, the decision matrix comprises the indicator values for each cluster center, as illustrated in Table 5. The following table presents the indicator weights that were calculated through the FGMM method. It is imperative to note that the values presented in this table have undergone a normalization process within the Minitab software.

The TOPSIS method was employed to extract the weight of the clusters according to the data in Table 5. To this end, Table 5 underwent normalization using the geometric normalization norm, the results of which are displayed in Table 6.

Subsequently, the weighted normal matrix was derived from

**Table 4.** Calculation of weight of indicators by the FGMM method.

| Indicators | Geometric Mean using Eq. (2) | Normalization using Eq. (3) | Defuzzification using Eq. (4) | The global weight using Eq. (5) |
|---|---|---|---|---|
| P/E ratio | (1.995, 3.047, 3.979) | (0.180, 0.377, 0.728) | 0.428 | 0.365 |
| Volume | (0.231, 0.289, 0.423) | (0.021, 0.036, 0.077) | 0.045 | 0.038 |
| Liquidity rating | (1.089, 1.732, 2.365) | (0.099, 0.214, 0.433) | 0.248 | 0.212 |
| ROE | (0.423, 0.637, 1.089) | (0.038, 0.079, 0.199) | 0.105 | 0.090 |
| Coefficient $\beta$ | (1.038, 1.817, 2.365) | (0.118, 0.225, 0.433) | 0.258 | 0.221 |
| Return on capital | (0.423, 0.567, 0.833) | (0.038, 0.070, 0.152) | 0.087 | 0.074 |
| Sum | (5.468, 8.089, 11.054) | — | 1.172 | 1 |

**Table 5.** Decision matrix based on the characteristics of each cluster's center.

| The Center of the clusters | P/E | Volume | Return on capital | ROE | Coefficient $\beta$ | Liquidity rating |
|---|---|---|---|---|---|---|
| Cluster center 1 | -0.161 | 11.256 | -0.5 | -0.586 | 0.099 | -1.530 |
| Cluster center 2 | -0.131 | -0.129 | -0.198 | -0.067 | 0.523 | 1.071 |
| Cluster center 3 | -0.230 | 8.757 | -0.734 | 1.326 | 0.645 | -1.524 |
| Cluster center 4 | -0.112 | -0.089 | -0.364 | -0.605 | -0.975 | -0.480 |
| Cluster center 5 | 6.394 | -0.223 | -0.531 | -1.459 | -0.573 | -0.207 |
| Cluster center 6 | -0.192 | 3.883 | -0.404 | 0.391 | 0.517 | 0.878 |
| Cluster center 7 | -0.075 | -0.007 | -0.251 | -0.184 | 0.509 | -0.819 |
| Cluster center 8 | -0.171 | -0.276 | 1.334 | 1.345 | -0.406 | 0.529 |
| The weight of indicators | 0.365 | 0.038 | 0.074 | 0.090 | 0.221 | 0.212 |

**Table 6.** The normalized decision matrix.

| The Center of the clusters | P/E | Volume | Return on capital | ROE | Coefficient $\beta$ | Liquidity rating |
|---|---|---|---|---|---|---|
| Cluster center 1 | -0.025 | 0.761 | -0.277 | -0.228 | 0.061 | -0.548 |
| Cluster center 2 | -0.020 | -0.009 | -0.110 | -0.026 | 0.320 | 0.383 |
| Cluster center 3 | -0.036 | 0.592 | -0.407 | 0.516 | 0.395 | -0.545 |
| Cluster center 4 | -0.017 | -0.006 | -0.202 | -0.236 | -0.597 | -0.172 |
| Cluster center 5 | 0.998 | -0.015 | -0.295 | 0.568 | -0.351 | -0.074 |
| Cluster center 6 | -0.030 | 0.263 | -0.224 | 0.152 | 0.316 | 0.314 |
| Cluster center 7 | -0.012 | 0.000 | -0.139 | -0.072 | 0.312 | -0.293 |
| Cluster center 8 | -0.027 | -0.019 | 0.740 | 0.524 | -0.248 | 0.189 |
| The weights of indicators | 0.365 | 0.038 | 0.074 | 0.09 | 0.221 | 0.212 |

the product of the elements of matrix 6 and the weights of their corresponding indices. This derivation is presented in Table 7.

Table 7 also presents the ideal good and bad options. The ideal good (or bad) option is derived from the maximum (or minimum) values associated with profit (or cost) indicators, and the minimum (or maximum) values related to cost (or profit) indicators. It is noteworthy that, according to user input, the liquidity rating index was the sole category designated as a cost index in this study. The subsequent step involves calculating the distances of each cluster center to the good and bad ideal options, followed by the calculation of the relative closeness distance (R.C.) index. The results of these calculations are presented in Table 8.

### 4.3 The third phase evaluation and validation

This section is dedicated to the evaluation of the results and the presentation of the sensitivity analysis. According to the findings presented in Table 8, cluster 5 has been identified as the leading cluster based on the indicators that have been analyzed. This cluster consists of five companies, identified by the symbols Khazin, Abada, Plask, Kodma, and Ghagerji. Consequently, stakeholders are empowered to evaluate these companies and make informed decisions about their future strategies. The metric employed for the purpose of ranking the clusters was the closeness of each cluster to the ideal solution, as well as its distance from the nonideal solution. Furthermore, the distances between each pair of clusters can be obtained to assess the homogeneity of clusters based on the selected indicators. Table 9 presents the relevant distances.

As demonstrated in Table 9, the distance between clusters 4 and 7 is the minimum among all distances. Notwithstanding, these clusters are ranked as the eighth and third clusters. The rationale underlying this methodology is that the clusters are evaluated based on their proximity to the optimal solution and their deviation from the non-ideal solution. This evaluation is not contingent upon the distances between the clusters. The ranking obtained for the clusters based on the TOPSIS method is susceptible to variation due to the weights assigned to the indicators. In order to ascertain the extent to which this ranking is influenced by each indicator, a sensitivity analysis is conducted. To this end, the TOPSIS model was executed six times, with the removal of one indicator in each iteration. Subsequently, a hierarchical ranking is applied to the clusters based on the

**Table 7.** The weighted normalized decision matrix.

| The center of the clusters | P/E | Volume | Return on capital | ROE | Coefficient $\beta$ | Liquidity rating |
|---|---|---|---|---|---|---|
| Cluster center 1 | -0.009 | 0.029 | -0.021 | -0.021 | 0.013 | -0.116 |
| Cluster center 2 | -0.007 | 0.000 | -0.008 | -0.002 | 0.071 | 0.081 |
| Cluster center 3 | -0.013 | 0.023 | -0.030 | 0.046 | 0.087 | -0.116 |
| Cluster center 4 | -0.006 | 0.000 | -0.015 | -0.021 | -0.132 | -0.036 |
| Cluster center 5 | 0.364 | -0.001 | -0.022 | 0.051 | -0.078 | -0.016 |
| Cluster center 6 | -0.011 | 0.010 | -0.017 | 0.014 | 0.070 | 0.067 |
| Cluster center 7 | -0.004 | 0.000 | -0.010 | -0.006 | 0.069 | -0.062 |
| Cluster center 8 | -0.010 | -0.001 | 0.055 | 0.047 | -0.055 | 0.040 |
| Ideal good | 0.364 | 0.029 | 0.055 | 0.051 | 0.087 | -0.116 |
| Bad ideal | -0.013 | -0.001 | -0.030 | -0.021 | -0.132 | 0.081 |

**Table 8.** Ranking of the clusters based on the distance of their centers from the good and bad ideal solutions.

| Centers of clusters | The Distance from the bad ideal | The distance from the good ideal | Relative closeness index | Ranking of clusters |
|---|---|---|---|---|
| Cluster center 1 | 0.395 | 0.247 | 0.385 | 4 |
| Cluster center 2 | 0.430 | 0.205 | 0.323 | 6 |
| Cluster center 3 | 0.387 | 0.303 | 0.439 | 2 |
| Cluster center 4 | 0.450 | 0.119 | 0.209 | 8 |
| Cluster center 5 | 0.210 | 0.400 | 0.656 | 1 |
| Cluster center 6 | 0.426 | 0.206 | 0.326 | 5 |
| Cluster center 7 | 0.384 | 0.248 | 0.393 | 3 |
| Cluster center 8 | 0.431 | 0.140 | 0.245 | 7 |

indicator that has been removed. The results of this study are presented in Table 10.

Some changes in cluster ranking occur by removing each indicator and accordingly ranking the clusters. We report these changes in four categories:

- No changes: As Table 10 shows, the ranking of the clusters is not changed by removing some indicators like volume and ROE.
- Slight changes: By removing some indicators, a very slight change occurs in the cluster rankings. For example, by removing return on capital, only the rankings of clusters 4 and 8 are swapped.
- Many but negligible changes: By removing some indicators, like the coefficient $\beta$ and liquidity rating, although more changes are observed in the cluster rankings, they can be neglected because they are not very significant. For example, by removing the coefficient $\beta$, we observe these changes in the cluster rankings: changing the rank of Cluster 1 from 4 to 3, Cluster 2 from 6 to 8, Cluster 4 from 8 to 5, Cluster 5 from 5 to 7, Cluster 7 from 3 to 4, and Cluster 8 from 7 to 6. Although many changes are observed, there is no significant difference in the cluster rankings.
- Essential and significant changes: By removing some indicators, significant changes occur in the cluster rankings. For example, by removing the P/E ratio, the rank of cluster 5 is changed from 1 to 7. In other words, when the P/E ratio is removed, cluster 5, which was known as the best cluster, is known as one of the worst clusters. This is for two reasons: first, the P/E ratio for cluster 5 is significantly higher than other clusters, and second, the experts assigned a weight to this indicator (0.365) much higher than the weights of other indicators. Therefore, if the weight of this indicator for experts is less important in similar studies, fundamental changes may be obtained in the rank of clusters.

Consequently, the findings of this sensitivity analysis indicate that the final cluster rankings of this study are stable and do not depend on a single indicator or its weight obtained by FAHP. The sole indicator that exhibited a substantial impact on the ranking of the clusters was the price-to-earnings (P/E) ratio.

## 5. Discussion and conclusion

The grouping of companies listed on the stock exchange is of crucial importance for financial analysis, forecasting company behavior, and developing portfolios based on these clusters. In most cases, this form of clustering is dependent upon an index that is specific to the industry in question. Nevertheless, it is noteworthy that companies across various industries may exhibit a greater degree of

**Table 9.** The distances between the centroids of clusters.

| Centers of clusters | Cluster center 1 | Cluster center 2 | Cluster center 3 | Cluster center 4 | Cluster center 5 | Cluster center 6 | Cluster center 7 | Cluster center 8 |
|---|---|---|---|---|---|---|---|---|
| Cluster center 1 | | 11.70 | 3.20 | 11.44 | 13.33 | 7.82 | 11.30 | 12.02 |
| Cluster center 2 | | | 9.37 | 2.23 | 6.89 | 4.04 | 1.89 | 2.35 |
| Cluster center 3 | | | | 9.26 | 11.64 | 5.52 | 8.93 | 9.55 |
| Cluster center 4 | | | | | 6.58 | 4.56 | 1.58 | 2.84 |
| Cluster center 5 | | | | | | 8.12 | 6.72 | 7.41 |
| Cluster center 6 | | | | | | | 4.28 | 4.71 |
| Cluster center 7 | | | | | | | | 2.75 |
| Cluster center 8 | | | | | | | | |

**Table 10.** Ranking of the clusters without considering one of the indicators.

| Clusters | Ranking by considering all indicators | Ranking the clusters without considering .... | | | | | |
|---|---|---|---|---|---|---|---|
| | | P/E | Volume | Return on capital | ROE | Coefficient $\beta$ | Liquidity rating |
| Cluster 1 | 4 | 3 | 4 | 4 | 4 | 3 | 6 |
| Cluster 2 | 6 | 5 | 6 | 6 | 6 | 8 | 3 |
| Cluster 3 | 2 | 1 | 2 | 2 | 2 | 2 | 2 |
| Cluster 4 | 8 | 8 | 8 | 7 | 8 | 5 | 8 |
| Cluster 5 | 1 | 7 | 1 | 1 | 1 | 1 | 1 |
| Cluster 6 | 5 | 4 | 5 | 5 | 5 | 7 | 4 |
| Cluster 7 | 3 | 2 | 3 | 3 | 3 | 4 | 5 |
| Cluster 8 | 7 | 6 | 7 | 8 | 7 | 6 | 7 |

similarity. The extent of similarity is contingent upon the indicators employed for the purpose of clustering. For instance, smaller firms from different sectors might share more similarities in certain indicators compared to both small and large firms within the same sector. Consequently, this study has grouped companies that are registered on the Tehran Stock Exchange without taking into account their respective industries. Instead, the study has focused on indicators that have been identified by users. The generated clusters were then subjected to a ranking procedure that employed the FAHP-TOPSIS method. A potential shortcoming of this research is the set of indicators that were employed. The study employed six indicators to cluster these companies: The following variables are to be considered: price-to-earnings (P/E) ratio, trading volume, liquidity rating, return on equity (ROE), beta coefficient $\beta$, and return on capital. The TOPSIS method was employed to assess the relative merits of the clusters, with the liquidity rating serving as the primary cost criterion, and all other indicators considered as profit criteria.

An index such as P/E, which is regarded as a profit criterion in this research, may be regarded as a cost criterion in another study. The decision regarding the allocation of capital is contingent upon the investors' investment objectives, whether those objectives pertain to the pursuit of growth stocks or value stocks. Furthermore, a high price-to-earnings (P/E) ratio does not inherently signify the value of a stock. A comprehensive analysis of these entities is imperative to substantiate such assertions. These limitations can be examined and analyzed for other indicators as well, which can be a topic for future research. The investigation of the clustering of companies registered in Tehran's Stock Exchange by considering other indicators can be another topic for future research. Indeed, the selection of indicators for clustering was informed by the insights of subject matter experts. However, numerous indicators for clustering exist, and experts may lack the knowledge to effectively screen them. In such cases, a range of methodologies can be employed, including statistical tests and feature selection techniques, to screen and validate the relevance of indicators for the purpose of clustering. In addition, the results of this research provide a foundation for future studies to address the financial analysis of the obtained clusters or the formation of portfolios based on these clusters. It is important to note that the determination of the number of clusters in this study was based on the opinions of experts in the field. However, in the absence of a predefined number of clusters, these metrics can be utilized to determine the optimal number of clusters. This determination can be made using various methods, including the elbow method, information criterion, cross-validation, silhouette score, or gap statistic. This issue may warrant further investigation in subsequent studies.

**Authors contributions**
Authors have contributed equally in preparing and writing the manuscript.

**Availability of data and materials**
The data that support the findings of this study are available from the corresponding author upon reasonable request.

**Conflict of interests**
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] Y. Chen, X. Zhao, and J. Hao. A novel mopso-sode algorithm for solving three-objective sr-es-tr portfolio optimization problem. *Expert Systems with Applications*, 233:120742, 2023.

[2] T.-C. Fu, C.-P. Chung, and F.-L. Chung. Adopting genetic algorithms for technical analysis and portfolio management. *Computers & Mathematics with Applications*, 66(10):1743–1757, 2013.

[3] S. Abolmakarem, F. Abdi, K. Khalili-Damghani, and H. Didehkhani. Futuristic portfolio optimization problem: wavelet based long short-term memory. *Journal of Modelling in Management*, 19(2):523–555, 2024.

[4] K. Deep, K. P. Singh, M. L. Kansal, and C. Mohan. A fuzzy interactive approach for optimal portfolio management. *Opsearch*, 46: 69–88, 2009.

[5] A. Yurtsal, Y. Karaömer, and A. İ. Benzer. A swarm intelligence optimization algorithm for cryptocurrency portfolio optimization. *Anemon Muş Alparslan Üniversitesi Sosyal Bilimler Dergisi*, 10(1): 347–363, 2022.

[6] Y. Wang and T. Aste. Dynamic portfolio optimization with inverse covariance clustering. *Expert Systems with Applications*, 213:118739, 2023.

[7] M. A. Akbay, C. B. Kalayci, and O. Polat. A parallel variable neighborhood search algorithm with quadratic programming for cardinality constrained portfolio optimization. *Knowledge-Based Systems*, 198:105944, 2020.

[8] M. Amiri, I. R. Vanani, S. Hossein, and R. H. Agha. Development of fuzzy artificial intelligence and multi-objective planning model to optimize the portfolio of investment companies. *Journal Of Business Intelligence Management Studies*, 9(36):243–302, 2021.

[9] J. Behera, A. K. Pasayat, H. Behera, and P. Kumar. Prediction based mean-value-at-risk portfolio optimization using machine learning regression algorithms for multi-national stock markets. *Engineering Applications of Artificial Intelligence*, 120:105843, 2023.

[10] G. J. Alexander, A. M. Baptista, and S. Yan. Portfolio selection with mental accounts and estimation risk. *Journal of Empirical Finance*, 41:161–186, 2017.

[11] D. Witkowska, K. Kompa, and M. Staszak. Indicators for the efficient portfolio construction. the case of poland. *Procedia Computer Science*, 192:2022–2031, 2021.

[12] A. Sarlak, M. Mohammadtalebi, and B. Mohammadtalebi. The effect of liquidity and credit risk on the relationship between business activities and fluctuations in the price of all companies listed on the tehran stock exchange. *Advances in Mathematical Finance and Applications*, 2(1):69–81, 2017.

[13] M. Alifamian, A. Eshaghzade, and A. Maleknia. Stock liquidity and return predictability; is there a connection?(evidence from an emerging market). *Advances in Mathematical Finance and Applications*, 8 (3):985–1006, 2023.

[14] M. Barbati, J. R. Figueira, S. Greco, A. Ishizaka, and S. Panaro. A multiple criteria methodology for priority based portfolio selection. *Socio-Economic Planning Sciences*, 88:101595, 2023.

[15] J. Puerto, M. Rodríguez-Madrena, and A. Scozzari. Clustering and portfolio selection problems: A unified framework. *Computers & Operations Research*, 117:104891, 2020.

[16] F. Jabbari-Moghadam, F. Hosseinzadeh Lotfi, M. Rostamy-Malkhalifeh, M. Sanei, and B. Rahmani-Parchkolaei. Computing the efficiency of bank branches with financial indexes, an application of data envelopment analysis (dea) and big data. *Advances in Mathematical Finance and Applications*, 9(3):817–839, 2023.

[17] M. E. Porter and C. Strategy. Techniques for analyzing industries and competitors. *Competitive Strategy*, page 1, 1980.

[18] V. Domazetoski, M. Rizinski, D. Trajanov, and L. Kocarev. Sensitivity analysis of shock distributions in the world economy. *Plos One*, 18(10):e0293524, 2023.

[19] S. Bhojraj, C. M. C. Lee, and D. K. Oler. What's my line? a comparison of industry classification schemes for capital market research. *Journal of Accounting Research*, 41(5):745–774, 2003.

[20] L. K. C. Chan, J. Lakonishok, and B. Swaminathan. Industry classifications and return comovement. *Financial Analysts Journal*, 63(6): 56–70, 2007.

[21] R. L. Phillips and R. Ormsby. Industry classification schemes: An analysis and review. *Journal of Business & Finance Librarianship*, 21(1):1–25, 2016.

[22] A. K. Birjandi, S. Dehmolaee, R. Sheikh, and S. S. Sana. Analysis and classification of companies on tehran stock exchange with incomplete information. *RAIRO-Operations Research*, 55:S2709–S2726, 2021.

[23] D. Farid, A. Dehghani Firizabadi, D. Andalib Ardakani, and H. Mirzaei. Analysis of factors affecting the selection of stock portfolio using the fuzzy logarithmic preference planning approach. *Farda Management Journal*, 66, 2021.

[24] P. Koratamaddi, K. Wadhwani, M. Gupta, and S. G. Sanjeevi. Market sentiment-aware deep reinforcement learning approach for stock portfolio allocation. *Engineering Science and Technology, an International Journal*, 24(4):848–859, 2021.

[25] F. Omidi, L. Torkzadeh, and K. Nouri. Mean-avar-skewness-kurtosis optimization portfolio selection model in uncertain environments. *Advances in Mathematical Finance and Applications*, 1:201–217, 2024.

[26] R. Darabi and M. Baghban. Application of clayton copula in portfolio optimization and its comparison with markowitz mean-variance analysis. *Advances in Mathematical Finance and Applications*, 1: 33–51, 2018.

[27] M. Zanjirdar. Overview of portfolio optimization models. *Advances in Mathematical Finance and Applications*, 1:419–435, 2020.

[28] J. Han, J. Pei, and H. Tong. Data mining: concepts and techniques. *Morgan kaufmann*, 2022.

[29] R. Maronna. Charu c. aggarwal and chandan k. reddy (eds.): Data clustering: algorithms and applications. *Statistical Papers*, 57(2): 565, 2016. URL https://doi.org/10.1007/s00362-015-0661-7.

[30] Q. Song, A. Liu, and S. Y. Yang. Stock portfolio selection using learning-to-rank algorithms with news sentiment. *Neurocomputing*, 264:20–28, 2017.

[31] J. Liu. Portfolio selection in stochastic environments. *The Review of Financial Studies*, 20(1):1–39, 2007.

[32] S. Kumari, N. Patil, P. Nankar, and M. Kulkarni. Cuda parallel computing framework for stock market prediction using k-means clustering. *2020 International Conference on Smart Electronics and Communication (ICOSEC)*, pages 467–473, 2020.

[33] P. D'Urso, L. De Giovanni, and R. Massari. Trimmed fuzzy clustering of financial time series based on dynamic time warping. *Annals of Operations Research*, 299(1):1379–1395, 2021.

[34] S. T. Tayalı. A novel backtesting methodology for clustering in mean-variance portfolio optimization. *Knowledge-Based Systems*, 209:106454, 2020.

[35] D. León, A. Aragón, J. Sandoval, G. Hernández, A. Arévalo, and J. Niño. Clustering algorithms for risk-adjusted portfolio construction. *Procedia Computer Science*, 108:1334–1343, 2017.

[36] H. Arman, A. Hadi-Vencheh, A. Arman, and A. Moslehi. Revisiting the approximated weight extraction methods in fuzzy analytic hierarchy process. *International Journal of Intelligent Systems*, 36(4): 1644–1667, 2021.