International Journal of Research in Industrial Engineering

www.riejournal.com

Int. J. Res. Ind. Eng. Vol. x, No. x (xx) x–x.

Paper Type: Original Article

The Optimization Model for Allocating Reward to Employees using GAHP and Cluster Analysis

Mehdi Ajalli*

Assistant Professor, Department of Industrial Management, Faculty of Management and Accounting, Bu-Ali Sina University, Hamedan, Iran; m.ajalli@basu.ac.ir

Abstract

Classification is one of the important tasks in any work and field. Cluster analysis (CA) is one of the most important classification methods. CA is one of the widely used methods in many scientific fields. Clustering is one of the most popular data mining techniques with many applications in industry. Especially, in the field of human resources management, the use of predefined rules is used to determine the performance and division of employees. The main goal of the current research is to design a suitable model for allocating rewards to employees by using the combined approach of the Group Analytical Hierarchy Process (GAHP) and CA. The research method is practical in terms of purpose and descriptive-survey in terms of data collection. For this purpose, first, by designing and distributing a comparative questionnaire of indicators and completing them by the experts of Shahid Fakuri Industries' component manufacturing unit, and by using the group hierarchical analysis process model with Expert Choice software, the weight of the effective indicators in employee evaluation was calculated, then the values of the indicators for 29 employees with using the formula of the normalization function in the Excel software, it is standardized and the weight of the indicators is multiplied by the standard values of the data, and then the distance matrix and the optimal number of clusters are calculated through the Machaon software, and finally, using the discriminative clustering approach and using the K-means method, data clustering was done with SPSS and Makaon software and a suitable model was presented for allocating rewards to the workers of the parts making unit.

Keywords: reward, employees, CA, GAHP, discriminative clustering, K-means method.

Corresponding Author:

1. Introduction

The turbulent and competitive world governing the business space of organizations has turned the human force component into an influential factor and a potential opportunity in maintaining and improving organizational resources and lasting stability of organizations; Thus, having quality human resources with skills appropriate to the needs of the organization is of particular importance, and the growth and development of organizations, which are important pillars of all-round social and economic development, requires the presence of developed people in organizations. Performance evaluation is a tool that helps organizations and employees in achieving organizational goals. If this tool is well designed and used correctly, it will be a suitable tool for encouraging, training and improving and sometimes correcting employees. In today's era, the tremendous developments in management knowledge have made the existence of performance evaluation system inevitable; in such a way that the lack of evaluation system in different dimensions of the organization, including evaluation in the use of resources, facilities, goals and strategies, managers and employees consider it as one of the symptoms of the organization's illness. The human element is considered the only factor in gaining a sustainable competitive advantage. One of the most important tasks of human resources management is to evaluate the performance of employees, and creating a suitable performance evaluation system provides a clear picture of the strengths and weaknesses of the organization and employees for the development of the organization and organizational growth. Most researches have not provided a comprehensive picture of performance evaluation indicators and in determining performance evaluation indicators, either only financial indicators have been considered or finally one or more qualitative indicators have been examined. One of the new methods in evaluating the performance of employees of organizations is to use the cluster analysis approach. Clustering is the grouping of similar objects using object data (Serra, 2017). One of the most important data mining methods (Reddy et al., 2013), clustering is among the most basic ways of data analysis with complete applications (Mousavian et al., 2022). In other words, in cluster analysis, it is tried so that the observations in each cluster (group) have the greatest similarity in terms of the desired variables, and the observations of each group have the greatest distance from the observations of other groups. For researchers and users, clustering and in its more general form, that is, segmentation, is not the end goal, but the beginning of other tasks. For example, in marketing, customers are first segmented based on different indicators (variables). Then, the behavior of each department is identified and planned for more appropriate and specialized service to each department. Cluster analysis is a branch of multivariate statistical analysis and unsupervised training in pattern recognition. This method is used to classify similar groups from a data set into similar clusters and dissimilar groups into dissimilar clusters. Classification of similar objects into multiple groups is one of the most important human activities. In everyday life, this is part of the learning process: the child learns how to distinguish between dogs and cats, between tables and chairs, between men and women, using mental schemas. This article shows why cluster analysis is often considered as a branch of pattern recognition and artificial intelligence. Classification has always played an essential role in science. In the past, clustering was based on visual-mental methods that were based on the researcher's judgment and perception. About 40 years ago, scientists in various sciences began to develop systematic methods for grouping data with the development of computers, cluster analysis grew rapidly.

Although the use of cluster analysis approach is widely developing, not much research has been done in the field of employee performance and awarding rewards to employees with better performance. In this research, by using the combined approach of group hierarchy analysis and cluster analysis, a suitable model for allocating bonuses to the workers of component manufacturing industries has been presented.

In a research, Serra (2017) presented an evaluation method to evaluate the performance of employees using cluster analysis. In this research, the criteria of call centers are considered in evaluating the performance of customer representatives. Cluster analysis and multidimensional scaling techniques are used to classify the performance criteria. The aim of this study is to classify customer representatives based on similar functional characteristics using cluster analysis. Based on the results, it was determined that the best variables for evaluating employee performance are the number of calls and sales values. In the cluster analysis, a clear clustering occurred especially according to the number of contacts. In multidimensional scaling results, it's easier to see which customer reps differ by other performance metrics, such as sales, sales offers, and cancellations. Employee performance evaluation is very important for both managers and employee motivation. Thus, a fair evaluation can be done using cluster analysis to evaluate employee performance.

Caryl (2019) conducted a research titled increasing organizational performance through employee training and development using k-Means cluster analysis. In this research, the raw data of 100 employees were evaluated and profiled to determine the appropriate training needed by identifying skill gaps. This cluster analysis was able to perform distinct classifications to determine changes in employee training needs to provide valuable insights for HR and organizational decision makers to understand how employee behavior and needs are changing. This allows companies to be in a better position to respond positively and provide appropriate training that meets the needs of individual employees by monitoring employee performance and identifying problem areas in a timely and accurate manner, unlike traditional training approaches used by many companies. This fast and efficient algorithm allows for a direct and objective talent list that will be the basis for future hiring, promotion and retraining decisions.

Bolisettyet al. (2020) in a study presented a performance evaluation model using unsupervised K-Means clustering. This paper seeks to implement an intelligent performance evaluation model that overcomes these factors and irregularities by using machine learning algorithms. The current scenario of performance appraisal involves more manual work in appraisal which is not accurate. This includes various aspects where the appraisal will only be favorable for a certain group of employees based on certain factors, but by using machine learning algorithms, the accuracy can be improved and increased to ensure that the employees receive the appraisal without any bias. A fair assessment that takes into account all factors is required. In this research, powerful machine learning systems have been used to build a model for employee performance classification. Also, clustering techniques were used to implement a self-learning model, that is, unsupervised learning models. In the end, a series of these techniques were used to increase accuracy.

Quang et al. (2021) have studied fuzzy clustering methods to understand employee performance in a research. In this article, the use of fuzzy clustering algorithms in the problem of human resource management is discussed. An approach is evaluated using a real-world HR dataset collected from a factory in Vietnam over ten years. According to the experimental results, the proposed approach has great potential in improving the understanding of employees' performance.

Mousavian et al., (2022) evelops an innovative auxiliary system for automatic labeling of numerical data by providing a hybrid clustering algorithm of K-means and partition around medoids (PAM) methods to identify organizational productive employees and to divide them into different productivity levels. The model is evaluated by calculating the differences between actual and labeled values (93% labeling accuracy) and an innovative criterion for image processing of the final clusters using the singular value decomposition (SVD) algorithm. Ultimately, the results of the algorithm determine four labels of middle and good productive employees who leave the organization and excellent and weak productive employees who stay in the organization; according to each cluster, policies are adopted for their retaining, productivity improvement, and replacement.

Cubukcu & Cantekin (2022) used a combined Fuzzy-AHP and TOPSIS decision model for selecting the best firewall alternative. This research offers a new solution related to a decision making problem that has started to gain more importance with the current digitalization process due to Covid-19 pandemic.

Ajalli et al., (2022) evaluated performance indicators in humanitarian logistics using path analysis, fuzzy DEMATEL and SWARA. The final finding of relationship analysis showed that "donation to delivery time" is the most influential indicator in terms of influencing other indicators. Such, results indicates the extraction of the fourth functional index i.e. "evaluation accuracy includes: speed and accuracy of committed donation and relief items delivered to stakeholders and how to assess the needs of stakeholders by employees" with the highest weight in the first rank as the most important functional indicator of humanitarian logistics and the second functional index i.e. "donation time includes "The delivery time of relief items in the country of destination after a donation and the collective remembrance of the donation" is important in the last rank.

Shakerian et al., (2023) used AHP and VIKOR method for selection of the best contractor of an executive project in the dairy industry. The findings state that contractor C received the first rank, contractor B the second rank and contractor A the third rank, and finally the first rank contractor was approved by the people present in the project to start and implement executive activities.

Chiniforooshan & Marinkovica (2023) used a hybrid particle swarm optimization algorithm for single machine scheduling with sequence-dependent setup times and learning effect. In order to test the applicability of the proposed algorithm to solve large-sized problems, 120 instances are generated, and the results are compared with a Random Key Genetic Algorithm (RKGA). The results show the effectiveness of the proposed model and algorithm.

Ganske and Carbon (2023) develop and establish a common ground to communicate and coordinate joint work efforts, which can mutually benefit and create synergies. The present article conceptualizes Effective internal Cluster Communication and Place Leadership as determinants for successful cluster developments. Despite the multiplicity of actors and sometimes even competing interest groups, Effective Place Leader Cluster Communication (EPLCC) enables clusters to inspire common, cooperative, collaborative and synergetic ways of working together. This is key to cluster development, successful and goal-directed cluster operation, and a sustainable operation of the cluster.

Abdel et al. (2024) in a study used MCDM methodology for selecting optimal charcoal company. This study used the Combinative Distance Assessment (CODAS) method as an MCDM method to rank the alternatives and use the best one. This study used nine criteria and twenty alternatives. The requirements are divided into positive and negative criteria to compete for the positive and negative ideal solution using the CODAS method. The criteria weights are computed. The rank of alternatives is checked by using the sensitivity analysis. The results show the rank of other options is stable in different cases.

The main goal of the current research is to design a suitable model for allocating rewards to employees of component manufacturing industries in Iran by using the combined approach of the GAHP and cluster analysis. The component manufacturing industry is a producer of all kinds of metal parts, which, by establishing quality, safety and occupational health management systems based on ISO9001:2008 and OHSAS18001:2007 standards, strives in line with its strategic goals by thinking of continuous and effective improvement in quality, safety and occupational health performance. Currently, this industry is one of the leading component manufacturing companies in the country with efficient manpower and well-equipped, advanced workshops, as well as various production processes, including machining, thermal operations, forging, casting, molding, etc., which shows its ability and capability. The company aims to satisfy the Islamic society of Iran. The main problem of the mentioned industry is the lack of a proper system for allocating rewards to employees based on performance. Considering the desire of the managers of this industry to conduct a research regarding the design and presentation of a suitable model for allocating rewards to employees, the researcher found that through interviews with industry experts and extracting the weight of indicators and using the cluster analysis approach, 29 of the employees in question The industry is categorized into distinct clusters and their rewards are allocated based on the extractive cluster and actual performance. For this purpose, the basic questions of the research are raised as follows:

1. What are the key performance indicators effective in allocating bonuses to industry employees?

2. What is the weight (importance) of the extracted indicators?

3. How is the clustering of industry workers based on the performance and weight of the mentioned indicators?

2. Methods

Optimization applications abound in many areas of science and engineering (Khalifa et al., 2023). Using MCDM procedures, the best alternatives are chosen from a collection of options after being evaluated according to several criteria (Abdel et al., 2024). As described, in order to optimize the allocation of reward to employees, MCDM optimization and clustering techniques have been used. Since most of the surveyed employees were married, this index did not have much effect on the clustering of the research data and was excluded from the study. In the process of clustering, after collecting data and standardizing them, the distance (dissimilarity) or similarity between objects should be calculated. The type of measurement scales is effective in calculating the distance or similarity. There are also different methods for calculating distance and similarity. In this section, an attempt has been made to present and review these methods. The $n \times n$ matrices whose values are the amount of distance (dissimilarity) are called distance matrix, and the matrix whose values are the similarity coefficient between objects is called similarity matrix. Distance and similarity matrices are also called proximity matrix or similarity matrix. Clustering is based on the proximity (similarity) matrix. If the proximity matrix, which we denote by $\mathrm{M_s}$ or $\mathrm{M_d}$ (where the index d represents the distance and the index s represents the similarity) is calculated based on the distance, its diameter is zero and the upper and lower values of the diameter are the same If it is calculated based on the similarity coefficient, the diameter of the matrix is one and the upper and lower values of the diameter are the same. In order to be symmetrical, sometimes only the upper or lower diameter values are written. These two matrices are shown below:

$$M_d = \begin{bmatrix} 0 & d_{12} & d_{13} & d_{1n} \\ d_{21} & 0 & d_{23} & d_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ d_{n1} & d_{n2} & \dots & 0 \end{bmatrix}$$

$$M_s = \begin{bmatrix} 1 & S_{12} & S_{13} & S_{1n} \\ S_{21} & 1 & S_{23} & S_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ S_{n1} & S_{n2} & \dots & 1 \end{bmatrix}$$

Deterministic clustering methods are generally divided into two categories: hierarchical and discriminative. In this research, separation methods have been used to cluster employees. Discrete clustering methods are very efficient for large problems. Big problems mean problems with a large number of objects, or a large number of indicators, or both. Separation methods are also called centripetal and non-hierarchical methods. In these methods, the number of clusters is already known. Its purpose is to determine which cluster each object belongs to. In this method, the error function is defined that we seek to minimize. In centrist methods, it is assumed that the clusters are convex shapes and the center of the cluster is a good representative for that cluster. Therefore, centripetal methods are not a suitable option for finding clusters with optional (non-convex) shapes. In this research, the K-means method is used as one of the separation methods used for clustering.

2.1. GAHP

Analytical hierarchy process (AHP) has been widely used for decision-making problems and successfully applied to many practical problems (Saaty, 1982, 1995, 2008). AHP is a well known MCDM technique to rank the decision alternatives. In AHP, you develop a hierarchy starting from upper level criteria and go one level down with each sub criteria and at the bottom level, you sort the alternatives (Cubukcu & Cantekin, 2022). The AHP by analyzing difficult and complex issues turns them into a simple form and solves them. This method has found many applications in economic and social issues and in recent years it has also been used in management affairs. In the AHP, the elements of each level are compared to their corresponding element at a higher level in pairs and their weight is calculated. We call these weights relative weights. Then by combining the relative weights, the final weight of each option is determined, which we call the absolute weight. In these comparisons, decision makers will use verbal judgments, so that if element i is compared with element j, the decision maker will say that the importance of i over j is one of the following situations:

- Completely preferred or completely more important or completely more desirable (9);
- Very strong preference or importance or desirability (7);
- Strong preference or importance or desirability (5;
- Slightly preferred or slightly more important or slightly more desirable (3);
- Same preference or importance or desirability (1).

These judgments have been converted by a clock into quantitative values between 1 and 9, which are specified in Table 1:

Table 1: Quantification of judgments

Preferences (oral judgments)		Numerical value
Extremely Prefered	ompletely preferred or completely more important or completely more desirable	9
Prefered Very Strongly	very strong preference or importance or desirability	7
Prefered Strongly	strong preference or importance or desirability	5
Moderately Prefered	slightly preferred or slightly more important or slightly more desirable	3
Equally Prefered	same preference or importance or desirability	1
-	Preferences between intervals	2,4,6,8

Many managers and experts are involved in many decision-making issues. The easiest way in such cases is to hold decision-making meetings and reach consensus among experts and managers. In short, the Group AHP is a method to perform an aggregate pairwise comparisons based on the viewpoint of a group of experts to improve decision making.

In order to calculate the weight of each option from the comparison matrix, several methods have been proposed, which are:

1. Ordinary least squares method;

2. Logarithmic least squares method;

3. Special vector method;

4. Approximate methods the arithmetic mean method for calculating relative weights has the following steps:

The first step: We add the values of each column together.

The second step: Each element in the pairwise comparison matrix is divided by the sum of its own column so that the pairwise comparison matrix is normalized.

The third step: We calculate the average value of the elements in each row of the normalized matrix. These average values are an estimate of the desired weights. If we want to use the opinions of several experts to calculate the relative weights (group decision), we must first sum up the K (number of experts) of the pairwise comparison matrix and arrive at a matrix that includes the opinions of all the experts. Each of the elements of the group pairwise comparison matrix is the geometric mean K of its corresponding number. Geometric mean of numbers X_1, X_2, … , X_K are defined as follows: $X = (X_1 \,.\, X_2 \ldots X_K)^{\wedge(1/K)}$

2.2. K-means method

The k-means method is the most practical data clustering method. This method was first presented by (Macqueen, 1967). The number of clusters in this method is fixed and predetermined. This method was designed to cluster data that are numerical (quantitative) and the cluster has a center called "mean". In this method, first, the objects are randomly divided into k clusters. In the next step, the distance of each object from the center of its cluster is calculated. If the distance of the desired object is greater than the average of its cluster and is closer to another cluster, this object is assigned to the cluster that is closer. This work is repeated until the error function is minimized, or the members of the clusters do not change. A clustering algorithm for machine learning enables organizations to be intuitive by making data-driven decisions and by ensuring effective human resource policies and various value-added interventions to an organization (Caryl, 2019).

If D is a data set with n objects, and C_1، C_2 و ... ،C_k represent k separate clusters of D, then the error function (EF) is defined as the sum of the distances of each object from the center of its own cluster (relation 3):

$$\mathrm{EF} = \sum_{i=1}^{k} \sum_{X \in C_i} d(X, \mu(C_i)) \tag{3}$$

Where μ indicates the center (mean) of the cluster, and $d(X, \mu(C_i))$ is the distance of each object from its center. To minimize it, centripetal clustering problems can be viewed as optimization problems. In this type of clustering, there is an error objective function that we seek to minimize, and limitations such as: a) the number of clusters is predetermined and cannot be increased or decreased, and b) the number the members of none of the clusters can be zero. In k-means clustering, steps are taken as follows:

The initial step: separating the primary data into k clusters arbitrarily;

Iterative step: a) calculating the distance of each objects from its center, b) calculating the error function;

Improvement step: moving the member with the largest distance from the center of its own cluster to the cluster with the smallest distance from it.

Stop command: not changing the members of the clusters or not reducing the value of the error function.

3. Research findings

In this research, all the employees related to one of the parts industry have been selected as samples for investigation. This component manufacturing industry has 6 key parts, and the current research was conducted based on the interest and sensitivity of the industry managers in structuring a suitable quantitative model for the purpose of evaluating and allocating rewards to employees. According to the good results of the proposed model in improving the status of reward allocation to employees of the selected unit, the researchers intend to evaluate the employees of other units of that industry in future researches and provide a final report on the status of reward allocation to employees. For this purpose, the information of the employees of the selected unit (job position; type of degree; work experience; hard work and marital status) was extracted from the administrative and human resources department of the industry. For this purpose, the information of the employees of the selected unit was extracted from the administrative and human resources department of the industry. In fact, the 29 employees studied in the selected industry unit are representatives of 6 key industry units that were randomly selected from among the units.

Next, the data of the indicators were collected for 29 employees. Considering that the measurement scale of the indicators was not the same, therefore, in the first step, data standardization should be done.

3.1. Data standardization

By dividing the data by the sum of the corresponding row (relation 2), the data are normalized:

$$z_{ij} = \frac{x_{ij}}{\sum_{j=1}^{d} x_{ij}} \quad (2)$$

It is clear that in this method, the sum of the normalized values will be equal to one. Table 1 shows the raw values of the criteria and Table 2 shows the normalized data values of the 4 indicators related to the research:

Table 2: Raw values for 4 criteria

Indicator		Job position	Type of degree	Work experience	Hard work	Marital status
Employee	1	45	25	17	3	married
Employee	2	90	45	25	3	married
Employee	3	105	45	6	3	married
Employee	4	85	45	4	3	married
Employee	5	65	25	14	3	married
Employee	6	60	35	14	3	married
Employee	7	65	35	14	3	married
Employee	8	65	35	14	3	married
Employee	9	60	35	14	3	married
Employee	10	60	25	12	3	married
Employee	11	60	25	12	3	married
Employee	12	35	18	12	3	married
Employee	13	95	45	7	3	married
Employee	14	65	35	17	3	married
Employee	15	95	25	20	5	married
Employee	16	65	35	17	3	married
Employee	17	95	55	15	3	married
Employee	18	85	45	6	3	married
Employee	19	85	45	9	3	married
Employee	20	130	55	8	3	married
Employee	21	130	35	22	3	married
Employee	22	85	45	5	3	married
Employee	23	115	45	14	3	married
Employee	24	45	25	20	7	married
Employee	25	45	25	18	6	married
Employee	26	45	14	14	6	married
Employee	27	45	13	13	6	married
Employee	28	51	14	14	6	married
Employee	29	60	12	12	5	married
Total		2131	978	389	107	

Considering that all the employees of the studied industry are married, therefore, the status of this ceriterion has no effect on the difference in remuneration between employees, and it is removed from the comparison and analysis.

Normalized values related to 4 criteria are presented in Table 3:

Table 3: Normalized values related to 4 indicators

Indicator		Job position	Type of degree	Work experience	Hard work
Employee	1	0.021	0.026	0.044	0.028
Employee	2	0.042	0.046	0.064	0.028
Employee	3	0.049	0.046	0.015	0.028
Employee	4	0.040	0.046	0.010	0.028
Employee	5	0.031	0.026	0.036	0.028
Employee	6	0.028	0.036	0.036	0.028
Employee	7	0.031	0.036	0.036	0.028
Employee	8	0.031	0.036	0.036	0.028
Employee	9	0.028	0.036	0.036	0.028
Employee	10	0.028	0.026	0.031	0.028
Employee	11	0.028	0.026	0.031	0.028
Employee	12	0.016	0.018	0.031	0.028
Employee	13	0.045	0.046	0.018	0.028
Employee	14	0.031	0.036	0.044	0.028
Employee	15	0.045	0.026	0.051	0.047
Employee	16	0.031	0.036	0.044	0.028
Employee	17	0.045	0.056	0.039	0.028
Employee	18	0.040	0.046	0.015	0.028
Employee	19	0.040	0.046	0.023	0.028
Employee	20	0.061	0.056	0.021	0.028
Employee	21	0.061	0.036	0.057	0.028
Employee	22	0.040	0.046	0.013	0.028
Employee	23	0.054	0.046	0.036	0.028
Employee	24	0.021	0.026	0.051	0.065
Employee	25	0.021	0.026	0.046	0.056
Employee	26	0.021	0.010	0.036	0.056
Employee	27	0.021	0.021	0.033	0.056
Employee	28	0.024	0.024	0.036	0.056
Employee	29	0.028	0.028	0.031	0.047

3.2. GAHP and weight calculation

In this research, after designing and distributing the questionnaire for the paire comparison of the indicators and completing them by 20 experts and using GAHP and entering the information in the Expert Choice software, the weight of the indicators was calculated as shown in Table 4:

Table 4: Calculated weight of indicators with Expert Choice software

Indicator	Job position	Type of degree	Work experience	Hard work
Weight	0.481	0.229	0.142	0.148

3.3. Determining the optimal number of clusters

In real problems, on the one hand, the number of objects to be clustered and on the other hand, the number of indicators based on which clustering is done is many. Therefore, manual calculations will be difficult even for the simplest clustering methods. Therefore, software is used for this issue.

In this research, SPSS and Machaon software were used to determine the optimal number of clusters and cluster analysis. If we want to determine the number of clusters, it is necessary to increase and decrease the number of clusters and calculate different indicators each time. In the present research (clustering of 29 people based on four indicators: job position, degree type, work experience and hard work), we have used k-means method for clustering. We base the optimal number of clusters on 8 reliability indices (c index, Davies-Bouldin, Dunn, Goodman-Kruskal, profile, isolation, Jaccard and Round). For this purpose, we first cluster the data with k=2 using the k-means method, then we repeat the cluster validation process with the aforementioned 8 indicators. That is, we calculated the validity of clustering 8 different times for k=2, noted the result of each one, and write all of them in Table 5:

Table 5: Calculation of credit indices for two clusters

Indicator	C index	Davies-Bouldin	Dunn	Goodman-Kruskal	Profile	Isolation	Jaccard	Round
Value	0.072	0.936	1.819	0.776	0.513	0.869	0.498	0.498

Now we cluster the data using the k-means method with k=3. Then we recalculate the 8 credit indicators in order. The result is given in Table 6:

Table 6: Calculation of credit indices for three clusters

Indicator	C index	Davies-Bouldin	Dunn	Goodman-Kruskal	Profile	Isolation	Jaccard	Round
Value	0.065	1.408	0.935	0.787	0.389	0.649	0.448	0.448

In the following, the data is clustered with the k-means method with k=4 and k=5, and then 8 credibility indices are calculated in order. In summary, for better conclusions, the credibility results of four clusters are given in Table 7:

Table 7: Summary of the validation results of the four clusters

Indicator	k=2	K=3	K=4	K=5	The right number of clusters
c index	0.072	0.065	0.059	0.025	5
Davies-Bouldin	0.936	1.408	1.262	1.324	2
Dunn	1.819	0.935	0.974	1.028	2
Goodman-Kruskal	0.776	0.787	0.790	0.932	5
Profile	0.513	0.389	0.422	0.463	2
Isolation	0.869	0.649	0.659	0.660	2
Jaccard	0.498	0.448	0.382	0.209	2
Round	0.498	0.448	0.382	0.209	2

As shown in the table, the best number of clusters based on the c index is 5 clusters (the lower the value of this index is, the better), based on the Davies-Bouldin index, 2 clusters (the lower the value of this index is, the better), based on Dunn index, 2 clusters (the higher the value of this index, the better), based on the Goodman-Kruskal index, 5 clusters (the higher the value of this index, the better), based on the profile index, 2 clusters (the higher the value of this index is better) based on the isolation index, 2 clusters (the higher the value of this index is, the better), based on the Jaccard index, 2 clusters (the higher the value and closer to one, the better) and based on the Round index, 2 clusters (the value of each The more and closer to one, the better). Therefore, considering that 6 out of 8 investigated indicators know the number of 2 clusters better, it seems that 2 clusters are more suitable for these data. As a result, in this research, based on 2 clusters, 29 employees are clustered.

3.4. Employee clustering with Makaon software

The results of the clustering of 29 employees using Makaon software are shown in Table 8:

Table 8: The final result of employee clustering considering the number of 2 clusters using Machaon software

Employee		Cluster
Employee	1	0
Employee	2	1
Employee	3	1
Employee	4	1
Employee	5	0
Employee	6	0
Employee	7	0
Employee	8	0
Employee	9	0
Employee	10	0
Employee	11	0
Employee	12	0
Employee	13	1
Employee	14	0
Employee	15	1
Employee	16	0
Employee	17	1
Employee	18	1
Employee	19	1
Employee	20	1
Employee	21	1
Employee	22	1
Employee	23	1
Employee	24	0
Employee	25	0
Employee	26	0
Employee	27	0
Employee	28	0
Employee	29	0

In the method of clustering employees with Makaon software, 2 clusters have been proposed, and based on this the employees were placed in 2 clusters. The employees of the first cluster will be given the first level reward and the employees of the second cluster will be allocated the second level reward.

3.5. Employee clustering with SPSS software

The results of data clustering using SPSS software are as follows: Table 9 shows the first output, which are the initial centers of the clusters:

Table 9: Initial centers of clusters

	Cluster		
	1	2	3
J	.010	.019	.029
M	.002	.011	.013
S1	.005	.002	.003
S2	.008	.004	.004

Table 10 is the second output that shows what changes occurred in the centers of the clusters in each iteration:

Table 10: Changes in cluster centers

	Change in Cluster Centers		
Iteration	1	2	3
1	.004	.003	.003
2	.000	.000	.001
3	.000	.000	.000

Table 11 is the third output that shows the final centers of each cluster:

Table 11: Final Cluster Centers

	Cluster		
	1	2	3
J	.011	.017	.027
M	.005	.009	.011
S1	.005	.005	.005
S2	.007	.004	.004

Table 12 is the fourth output that specifies the number of members of each cluster, the number of valid and missing objects:

Table 12: The number of "members of each cluster", "valid objects" and "missing ones"

Cluster	**1**	10.000
	2	15.000
	3	4.000
Valid		29.000
Missing		.000

Table 13 is also the last output. The QCL_1 line indicates which cluster each object is a member of. The line QCL_2 also shows the distance of each object from the center of its cluster:

Table 13: Final clustering of employees

Employee	**1**	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**
QCL_1	1	2	3	2	2	2	2	2	2	1
QCL_2	0.00308	0.00537	0.00391	0.00428	0.00412	0.00372	0.00281	0.00281	0.00372	0.00404
Employee	**11**	**12**	**13**	**14**	**15**	**16**	**17**	**18**	**19**	**20**
QCL_1	1	1	2	2	2	2	2	2	2	3
QCL_2	0.00404	0.00432	0.00425	0.00311	0.00606	0.00311	0.00516	0.00349	0.00282	0.00336
Employee	**21**	**22**	**23**	**24**	**25**	**26**	**27**	**28**	**29**	
QCL_1	3	2	3	1	1	1	1	1	1	
QCL_2	0.00488	0.00349	0.00115	0.00428	0.00288	0.00359	0.00217	0.00345		

In the method of clustering employees with SPSS software, 3 clusters have been proposed, and based on this the employees were placed in 3 clusters. The employees of the first cluster are given the first level reward, the employees of the second cluster are given the second level reward, and the employees of the third cluster are allocated the third level reward.

4. Conclusions and suggestions

In some sciences, after clustering, scientists seek to see how these clusters are related to each other in order to extract "theories" from it. In this case, they look for a "communication bridge" between different sectors and clusters. In this research, a suitable model for allocating bonuses to the employees of Shahid Fakuri Industries component manufacturing unit has been presented. In this way, after designing and distributing the questionnaire for the two-way comparison of indicators and completing them by 20 experts, using the hierarchical analysis process approach and entering information into the Expert Choice software, the weight of the indicators was calculated. Based on the output of the calculated weight, the index of job position has the highest weight (importance) and the indicators of degree type, difficulty of work and work experience were placed in the next ranks in terms of importance. In the following, due to the fact that Makaon software knows the number of 2 clusters better with 6 validity indices out of 8 examined indices, 2 clusters were considered for data clustering. Therefore, in this research, based on 2 clusters, 29 employees were clustered using the K-means method. Further, clustering of 29 employees in 3 clusters was done through SPSS and Makaon software. The lack of familiarity of experts with the GAHP technique and cluster analysis was one of the main limitations of the research. Such, Hard access to experts due to not having enough time was one of the main and practical limitations of the present research. Because the real information of the employees was requested from the administrative and human resources unit of the industry, the sensitivity of this unit in providing the confidential information of the employees and obtaining the relevant permits from the industry protection unit caused a lot of time to be taken from the researchers and a lot of waiting to get the information.

In the next researches, it is possible to calculate the weight of indicators by using other modern multi-indicator decision-making approaches, such as the best-worst method (BWM), SWARA, etc. Also, by using the new approaches of multi-criteria decision-making such as ARAS, WASPAS, etc., it determined the ranking of employees in terms of performance and compared it with the cluster analysis method presented in this research.

References

1. Abdel Aal, S. I., J., S., & P. K., U. (2024). Selecting Optimal Charcoal Company using Multi-Criteria Decision Making Methodology. Multicriteria Algorithms with Applications, 3, 15-22.
2. Ajalli, M., Saberifard, N., & Zinati, B. (2022). Evaluation and ranking of performance indicators in humanitarian logistics using path analysis, fuzzy DEMATEL and SWARA. *Journal of Decisions and Operations Research*, *6*(Special Issue), 1-20.
3. Aldenderfer M.S. and Blashfield, R.K. (1984), Cluster analysis, Sage publications.
4. Bock, H. (1989), Probabilistic aspects in cluster analysis, In Optiz, O., Editor, Conceptual and Numerical Analysis of Data, 12-44. Agusburg, FRG: SPRINGER-Verlag.
5. Bolisetty Kanaka Durga Prasad; Bhaskar Choudhary; B Ankayarkanni (2020). Performance Evaluation Model using Unsupervised K-Means Clustering, International Conference on Communication and Signal Processing (ICCSP), 28-30 July, Chennai, India.
6. Caryl Jimenez (2019). Enhancing Organizational Performance through Employee Training and Development using k-Means Cluster Analysis, International Journal of Advanced Trends in Computer Science and Engineering 8(4):1576-1582.
7. Chiniforooshan, P., & Marinkovic, D. (2023). A hybrid particle swarm optimization algorithm for single machine scheduling with sequence-dependent setup times and learning effects. Computational algorithms and numerical dimensions, 2(2), 74-86.
8. Cubukcu, C., & Cantekin, C. (2022). Using a combined fuzzy-AHP and topsis decision model for selecting the best firewall alternative. Journal of fuzzy extension and applications, 3(3), 192-200.
9. Davies, D. and Bouldin, D. (1979), A cluster separation measure, IEEE Transactions on Pattern Analysis and Machine Intelligence, 1(2): 224-227.
10. Ganske Patricia and Carbon Claus-Christian (2023). Successful clusters through successful communication: why clusters should be managed by Place Leaders. Front. Commun. 8:1194103.
11. Gun, G.Ma, C. and Wu, J. (2007), Data clustering: theory, algorithms, and applications, Asa-Siam.
12. Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002a), Clustering validity methods: Part I. ACM SIGMOD Record, 31(2).
13. Halkidi, M., Batistakis, Y., and Vazirgiannis, M. (2002b), Clustering validity checking methods: Part II. ACM SIGMOD Record, 31(3).
14. Holman, E.W. (1985), Evolutionary and psychological effects in prerevolutionary classifications, Journal of Classification, 2:29-39.
15. Khalifa, H. A. E., Edalatpanah, S. A., & Bozanic, D. (2023). Enhanced a novel approach for smoothing data in modeling and decision-making problems under fuzziness. Computational algorithms and numerical dimensions, 2(3), 163-172.
16. Macqueen, J. (1967), Some methods for classification and analysis of multivariate observations, In Proceedings of the 5th Berkeley symposium on mathematical statistics and probability, volume 1, pages 281-297, Berkeley, CA: University of Califirnia.

17. Mousavian Anaraki Seyed Alireza, Haeri Abdorrahman, Moslehi Fateme (2022). Providing a Hybrid Clustering Method as an Auxiliary System in Automatic Labeling to Divide Employee into Different Levels of Productivity and Their Retention, Iranian Journal of Management Studies (IJMS), 15(2): 207-226.
18. Quang-Vinh Dang; Minh-Tuan Truong; Minh-Hoang Huynh (2021). 4th International Conference on Artificial Intelligence and Big Data (ICAIBD), 28-31 May, Chengdu, China.
19. Romesburg, H.C. (1984), Cluster analysis for researchers, Lulu Press.
20. Rousseuw, P.J. (1987) Silhoettes: a graphical aid to the interpretation and validation of cluster analysis, Journal of Computational and Applied Mathematics, 20, 53-65.
21. Rung-Wei P., Yuh-Yuan G., Miin-Shen Y. (2009), A new clustering approach using data envelopment analysis, European Journal of Operational Research,, 1999, PP. 276-284.
22. Saaty, (1995). Decision making for leaders. The analytical hierarchy process for decisions in a complex world. RWS publications, Pittsburgh.
23. Saaty, (2008). Decision making with the analytic hierarchy process. International Journal of Services Sciences 1(1), 83-98.
24. Saaty, T.L., (1982). Decision making for leaders. Lifetime learning publications. Wadsworth, Belmont.
25. Serra Çelik (2017). An Assessment Method for Employees Performance Evaluation Using Cluster Analysis, Proceedings of International Academic Conferences 5908119, International Institute of Social and Economic Sciences.
26. Shakerian Hagigi, S., Nezhad Heydari, Z., raeiszadeh, M., & Qayumi, M. (2023). Selection of the best contractor by AHP and VIKOR method, a case study of an executive project in the dairy industry. *Modern Research in Performance Evaluation*, *2*(2), 79-90.