Evaluating Employee Performance with an Improved Clustering Algorithm

Ci Fan
Zhengzhou University of Industrial Technology, Zhengzhou, Henan 451150, China
E-mail: f318to@163.com

This paper introduces performance appraisal management, analyzed its principles, proposes seven appraisal indicators, puts forward an improved K-means algorithm for the classification of appraisal results based on the K-means clustering algorithm and density parameters, and analyzes the performance of the method. The results of the UCI data set show that the accuracy rate of the improved K-means algorithm is significantly higher than that of the traditional K-means algorithm, and its highest accuracy rate is 91.27%. The classification of the appraisal results of 45 employees in Company A shows that the classification results obtained using the improved K-means algorithm are more in line with the actual situation and more reasonable and relevant than the traditional scoring method. The appraisal results of 1064 technical employees show that most of the employees have a good performance. The results of the experiment verify the effectiveness of the improved K-means algorithm in performance appraisal management. The algorithm can be further promoted and applied in practice.

Povzetek: Opisan je nov sistem za ocenjevanje kvalitete prispevkov sodelavcev s pomočjo gručenja.

1 Introduction

In order to make employees more motivated and responsible, it is a very important element to do a good performance appraisal management of employees [1]. With the development of society, people's life is also increasingly moving towards informationization and intelligence, and enterprise management also applies more technological tools [2]; therefore, performance appraisal management through a more intelligent way is of great significance to improve employee performance and increase the sense of belonging [3]. Most of the current performance appraisal management has a strong subjective nature and relies more on the subjective judgment of the appraiser, which cannot achieve objective and fair appraisal of employees; therefore, with the development of information technology, there are more and more intelligent methods applied in the performance appraisal management of company employees. Yousif et al. [4] conducted a study on the performance evaluation of faculty members in Sudanese universities and developed appropriate criteria. Then, relevant data were obtained through expert assessment. Faculty members were comprehensively evaluated through hierarchical analysis and a technique for order preference by similarity to an ideal solution (TOPSIS). Ditzian et al. [5] evaluated three supervisors using the Performance Diagnostic Checklist–Human Services (PDC-HS) and found that the PDC-HS-indicated intervention (graphed feedback) was effective in improving performance. Bird et al. [6] considered that a performance management approach combining individual competencies with expertise could create significant values within highly educated work forces. Karpenko et

al. [7] proposed an approach based on the overall key performance indicators (KPI) to analyze employee performance and found that through the use of KPIs, the process of operational activities can be better controlled and productivity can be improved. In this paper, an improved K-means algorithm was used to realize the performance assessment management of employees, and experimental analysis was carried out on Company A to verify the effectiveness of the method. The proposed method can realize a reliable performance management assessment, create a good working environment for employees, improve their motivation and promote better and faster development of the company.

2 Employee performance appraisal management indicators

Employee performance appraisal management is an important part of enterprise management [8], which is closely related to many tasks of enterprise management. For example, when conducting employee training, the results of the appraisal can reflect the technical level of employees, so training will be more targeted; when promoting positions, the results of the appraisal can help understand the ability, quality and attitude of employees; therefore, performance appraisal management has a certain effect on improving employees' motivation and dedication [9]. With the development of the economy and the increase of the importance of performance appraisal [10], the amount of employee performance information has become larger and larger, and the traditional methods are increasingly unable to meet the current needs of company management [11]. There are more and more new

methods that are applied for analysis in performance appraisal.

The performance appraisal management of employees generally follows the following principles.
(1) Openness principle: the development, criteria, process and result of the appraisal should be open.
(2) Objectivity principle: the criteria and process of the appraisal should be objective and avoid subjective emotions as much as possible.
(3) Identity principle: the same standard and scale are used to appraise employees.
(4) Institutional principle: performance appraisal management should be fixed as a system and strictly enforced.
(5) Feedback principle: the results of the appraisal must be fed back to the employees and relevant departments.
(6) Comprehensiveness principle: the content of the appraisal should be as comprehensive as possible.

In order to conduct performance appraisal, the indicators for appraisal need to be developed first. Through literature review and expert consultation, seven indicators are identified in this paper, which is as follows.
(1) Team spirit: able to obey the work schedule of the team.
(2) Creative spirit: the ability to come up with innovative ideas and approaches when faced with problems.
(3) Sense of responsibility: a high sense of responsibility to the company and work.
(4) Cost consciousness: able to use company resources effectively and save costs in the workplace.
(5) Working ability: Good grasp of relevant knowledge and skills, with high efficiency and high performance.
(6) Communication skills: smooth communication with all levels of departments and colleagues without conflicts.
(7) Moral quality: good conduct, decent behavior and good image.

After evaluating the scores of the above seven indicators, this paper uses an improved clustering algorithm to process and classify the indicator evaluation results into four categories: excellent, good, qualified, and unsatisfactory.

3 An improved clustering algorithm

The principle of a clustering algorithm is to divide the huge amount of data into different clusters according to certain similarities, so that objects with high similarity can be in the same class [12]. The clustering algorithm can analyze data without knowing the size and distribution [13], so it has good processing effects on the uncertainty of data [14]. The K-means algorithm is a commonly used classification-based algorithm with very common applications in the classification of data [15] and performs well in fields of medical detection [16] and image segmentation [17]. Its calculation steps are as follows.
(1) The number of clusters to be divided (K) is determined. According to the results of the performance assessment, K = 4 in this paper.

(2) For the data set $(x_1, x_2, x_3, x_4, x_5, x_6, x_7)$, seven initial clustering centers, $(C_1, C_2, C_3, C_4, C_5, C_6, C_7)$, are randomly selected. The distance from every object to the center is calculated, and then every object is divided into the closest class.
(3) The central point of every cluster is recalculated, and the formula is:

$$C'_i = \frac{1}{n_i} \sum_{x_j \in C_i} x_j,$$

where i = 1,2,3,4,5,6,7 and n_i refers to the number of objects in class C_i.

(4) If any $C_i = C'_i$ holds, the algorithm converges, and if not, the algorithm returns to the second step.
(5) The final clustering result is output.

The result of the K-means algorithm is easy to understand and stable; however, its result has some dependence on the selection of the initial clustering centers. To solve this problem, this paper improves the K-means algorithm by combining the density parameter of objects.

The density parameter of an object refers to the distance between an objective and the nearest Minpts-th (constant) object [18] centered on that object, denoted by ε. The greater the value, the lower the density of the region in which the object is located. In this paper, the parameter is improved: a new m-dist value is put forward, which is the mean value of the distance between an object and the nearest Minpts-th (constant) object. The flow of the improved K-means algorithm is as follows.
(1) For a given data set, all m-dist values are calculated based on the constant Minpts, and then the mean value, mean-dist, is calculated to obtain the set of density values, C.
(2) The object with the lowest density in set C is used as the first initial cluster center.
(3) The rest cluster centers are obtained in the same manner. Finally, seven initial clustering centers located in high-density areas are identified: $(C_1, C_2, C_3, C_4, C_5, C_6, C_7)$.
(4) Clustering is performed according to the K-means algorithm, and the result is output.

4 Example analysis

To verify the performance of the improved K-means algorithm, experiments were first conducted on the standard UCI data set [19], as shown in Table 1.

The clustering analysis of the above data set was performed using the traditional K-means algorithm and the improved K-means algorithm, respectively, and the experimental results are shown in Figure 1.

It was seen from Figure 1 that the accuracy of the traditional K-means algorithm ranged from 60% to 75%, with a minimum of 64.26% and a maximum of 74.51%, while the accuracy of the improved K-means algorithm was above 80%, with a minimum of 81.94% and a maximum of 91.27%, showing its reliability. The above results demonstrated that the accuracy of the algorithm was significantly enhanced after improvement; thus, it

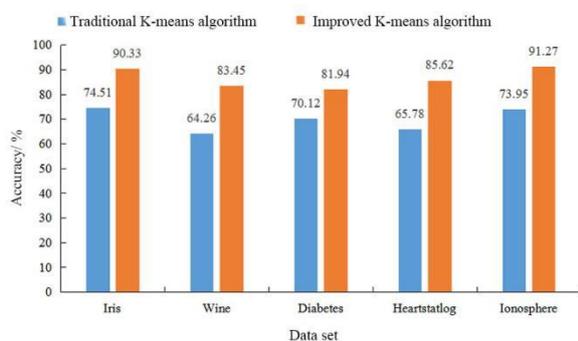| Databases	Instances	Number of classes (k)	Number of features (k)	Size of classes
Iris	150	3	4	505050
Wine	178	3	13	597148
Diabetes	768	2	8	268500
Heartstatlog	270	2	13	150120
Ionosphere	351	2	34	126225

Table 1: Experimental data set.

Figure 1Results of the experiment on the UCI data set.

would have better performance in performance appraisal management.

Then, to study the application effect of the improved K-means algorithm in employee performance appraisal management, Company A was taken as an example. The positions in the company included technical positions, management positions and scientific positions, and 15 employees were randomly selected from each of the three positions for performance management appraisal. The evaluators determined the score of every indicator according to the characteristics of different positions. The score of every position on every indicator is shown in Table 2. The total score of the appraisal was 70 points.

The indicators were scored by the appraisers, and the scoring results are shown in Table 3.

The assessment results were classified using the traditional scoring method and the improved K-means algorithm, respectively, and the classification criteria of the scoring method were: excellent for 45 points and above, good for 40-45 points, qualified for 35-40 points, and poor for 35 points and below. The final classification results of the two methods are shown in Table 4.

It was seen from Table 4 that the results obtained by the two classification methods were not exactly the same, and there were obvious differences. The results of technical positions were used for specific analysis. Firstly, the scores of different indicators for employees in technical positions are shown in Table 5.

	Technical positions	Management positions	Scientific research positions
Team spirit	12	10	12
Spirit of innovation	8	8	14
Awareness of responsibility	12	12	8
Cost awareness	8	8	10
Working ability	12	10	10
Communication skills	8	12	6
Moral quality	10	10	10
Total score	70	70	70

Table 2: Score of different positions (unit: point).

Number	Team spirit	Spirit of innovation	Awareness of responsibility	Cost awareness	Working ability	Communication skills	Moral quality	Total score
1	9	6	10	8	10	7	8	58
2	7	7	7	7	7	7	7	49
3	7	5	7	6	3	5	7	40
4	5	6	7	8	9	7	8	50
5	6	4	3	4	5	4	5	31
6	2	7	7	5	7	5	8	41
7	8	7	7	5	9	8	4	48
......	......	......	......	......	......	......	......	......
43	7	7	5	5	5	4	6	39
44	8	9	5	7	7	5	5	46
45	7	8	5	5	5	4	4	38

Table 3: Scores of every employee on different indicators.

It was seen from Table 4 that there were some differences between the assessment results of the two methods, and the data in Table 5 were combined to analyze the characteristics of different employees. The analysis results are as follows.

(1) Employee 7: Table 5 shows that Employee 7 had a total score of 48 points, i.e., excellent in the scoring method, but in terms of the specific indicators, his score in "moral quality" was only four points, which was one of the lowest scores among the 15 employees who participated in the assessment, while the scores of the remaining six indicators were all high,

		The scoring method	The improved K-means algorithm
Technical positions	Excellent	1, 2, 4, 7, 9.	1, 2, 4, 9
	Good	3, 6, 8, 12	8, 12, 15
	Qualified	14, 15	3, 7, 14
	Poor	5, 10, 11, 13	5, 6, 10, 13
Management positions	Excellent	16, 25, 28, 29	16, 25, 28
	Good	19, 21, 24, 30	18, 19, 24, 30
	Qualified	17, 18, 20, 26, 27	20, 21, 22, 26, 27, 29
	Poor	22, 23	17, 23
Scientific research positions	Excellent	31, 32, 37, 42, 44	31, 32, 44
	Good	34, 35, 38	34, 38
	Qualified	33, 40, 41, 43, 45	35, 37.
	Poor	36, 39	42, 36, 39

Table 2 Comparison of classification results.

indicating that his performance in moral quality was poor. Therefore, it was more reasonable to classify him as "qualified" in the improved K-means algorithm.

(2) Employee 3: Table 5 shows that the total score of Employee 3 was 40 points, i.e., good in the scoring method, but in terms of the specific indicators, his score in the "work ability" was only three points, indicating that his work ability was poor so that he was inappropriate for a technical position. Therefore, it was more reasonable to classify him as "qualified" in the improved K-means algorithm.

(3) Employee 6: Table 5 shows that Employee 6 had a total score of 41 points, but his score on "team spirit" was only two points, which indicated that his performance in team spirit was extremely poor and he might have disobedience and inability to cooperate well. Therefore, it was more reasonable to put him in the "poor" category in the improved K-means algorithm.

(4) Employee 15: Table 5 shows that the total score of Employee 15 was 39 points, i.e., qualified in the scoring method, but he was classified as good in the improved K-means algorithm. Specifically, the performance of Employee 15 was balanced in all indicators, all above five points, so it was more reasonable to classify him as "good".

The comparison and analysis of the appraisal results found that compared with the traditional scoring method, the improved K-means method obtained more practical results and provided a more comprehensive and detailed appraisal of employees, which can provide more reliable

Number	Team spirit	Spirit of innovation	Awareness of responsibility	Cost awareness	Working ability	Communication skills	Moral quality	Total score
1	9	6	10	8	10	7	8	58
2	7	7	7	7	7	7	7	49
3	7	5	7	6	3	5	7	40
4	5	6	7	8	9	7	8	50
5	6	4	3	4	5	4	5	31
6	2	7	7	5	7	5	8	41
7	8	7	7	5	9	8	4	48
8	6	5	6	5	7	5	7	41
9	6	6	8	6	7	6	8	47
10	7	4	2	4	7	4	6	34
11	5	4	5	4	5	4	4	31
12	7	6	7	6	7	5	6	44
13	4	4	5	4	5	4	4	30
14	6	5	3	6	8	5	6	39
15	6	5	6	6	6	5	5	39

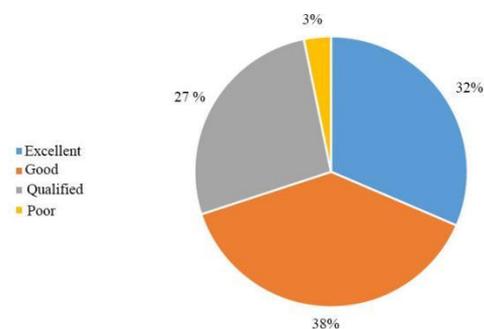Table 1 Score results of each index for technical positions.

Figure 2 Technical position assessment results.

and powerful support for the performance appraisal management of employees.

The performance appraisal of 1064 technical employees in Company A was conducted using the improved K-means method, and the results are shown in Figure 2.

It was seen from Figure 2 that the number of technical employees who were assessed as "good" was the largest, reaching 38%, followed by "excellent" employees, reaching 32%, and "qualified" employees, reaching 27%, and the number of employees who were assessed as "poor" was the lowest, 3%. According to the above results, most technical employees had a good performance, but there are also a small number of employees with poor performance. The analysis of the specific indicators suggested that the employees with poor performance either had an obvious lack of ability in a certain area or had poor overall capacity. For poorly performed employees, the company needs to carefully consider their salaries and promotions; for well-performed employees, the company should make appropriate encouragement to mobilize their enthusiasm.

The experimental analysis found that the improved K-means algorithm had favorable performance in performance appraisal management, which is conducive to improving the level of company management, promoting the efficiency and quality of management, establishing a reasonable human resource incentive mechanism, helping employees to discover their shortcomings, improve their working ability, and drive team development to promote company progress.

5   Conclusion

Based on the clustering algorithm, this paper studied the method of company employee performance appraisal management. An improved K-means algorithm was designed, and its effectiveness was verified through experiments on the UCI data set. Then, the actual employee performance data were classified. The comparison between the improved algorithm and the traditional scoring method showed that the improved algorithm had higher rationality and could be further promoted and applied in practice. This work makes some contributions to improve the performance appraisal management effect and promote the common progress of employees and companies.

References

[1]   Moradi T, Mehraban M A, Moeini M (2017). Comparison of the Perceptions of Managers and Nursing Staff Toward Performance Appraisal. Iranian journal of nursing and midwifery research, 22, pp. 128.

[2]   Najjar A, Amro B, Macedo M (2021). An intelligent decision support system for recruitment: resumes screening and applicants ranking. Informatica, 45, pp. 617-623.

[3]   Gomez-Mejia LR (2016). Increasing Productivity: Performance Appraisal and Reward Systems. Personnel Review, 2016, 19, pp. 21-26.
https://doi.org/10.1108/00483489010138759

[4]   Yousif M K, Shaout A (2016). Fuzzy logic computational model for performance evaluation of Sudanese Universities and academic staff. Journal of King Saud University - Computer and Information Sciences, 30, pp. 80-119.
https://doi.org/10.1016/j.jksuci.2016.08.002

[5]   Ditzian K, Wilder D A, King A, Tanz J (2015). An evaluation of the Performance Diagnostic Checklist-Human Services to assess an employee performance problem in a center-based autism treatment facility. Journal of Applied Behavior Analysis, 48, pp. 199.
https://doi.org/10.1002/jaba.171

[6]   Bird H (2015). Appraising clever people: lessons from introducing performance reviews for academics in a UK University. Industrial & Commercial Training, 47, pp. 81-85.
https://doi.org/10.1108/ICT-10-2014-0069

[7]   Karpenko A, Karpenko N, Shudrik O (2020). Development and implementation of a strategic personnel management system according to goals based on key performance indicators. Management and Entrepreneurship Trends of Development, 2, pp. 22-35.
https://doi.org/10.26661/2522-1566/2020-2/12-02

[8]   Uysal G (2016). Cognitive Placement Theory for Performance Appraisal: Talent Management and Individual Performance. Journal of Modern Accounting and Auditing, 012, pp. 60-63.
https://doi.org/10.17265/1548-6583/2016.01.005

[9]   Nair M S, Salleh R (2015). Linking Performance Appraisal Justice, Trust, and Employee Engagement: A Conceptual Framework. Procedia - Social and Behavioral Sciences, 211, pp. 1155-1162.
https://doi.org/10.1016/j.sbspro.2015.11.154

[10] Andersson G (2016). Data Recording in Performance Management: Trouble With the Logics. American Journal of Evaluation, 38, pp. 190-204.
https://doi.org/10.1177/1098214016681510

[11] Lonsdale A (2016). Performance Appraisal, Performance Management and Quality in Higher Education: Contradictions, Issues and Guiding Principles for the Future. Australian Journal of Education, 42, pp. 303-320.
https://doi.org/10.1177/000494419804200307

[12] Lv Y, Ma T, Tang M, Cao J, Tian Y, Al-Dhelaan A, Al-Rodhaan M (2016). An efficient and scalable density-based clustering algorithm for datasets with complex structures. Neurocomputing, 171, pp. 9-22.
https://doi.org/10.1016/j.neucom.2015.05.109

[13] Kumar Y, Sahoo G (2014). A chaotic charged system search approach for data clustering. Informatica, 38, pp. 249-261.

[14] Guo Y, Sengur A (2015). NCM: Neutrosophic c-means clustering algorithm. Pattern Recognition, 48, pp. 2710-2724.
https://doi.org/10.1016/j.patcog.2015.02.018

[15] Dhanachandra N, Manglem K, Chanu Y J (2015). Image Segmentation Using K -means Clustering Algorithm and Subtractive Clustering Algorithm. Procedia Computer Science, 54, pp. 764-771.

[16] Abaker AA, Saeed FA (2021). A Comparative analysis of machine learning algorithms to build a predictive model for detecting diabetes complications. Informatica, 45, pp. 117-125.

[17] Jiang HN (2018). Defect features recognition in 3d industrial ct images. Informatica, 42, pp. 477-482.

[18] Yuan F, Zhou ZY, Song X (2007). K-means Clustering Algorithm with Meliorated Initial Center. Computer Engineering, 33, pp. 65-66. https://doi.org/10.3969/j.issn.1000-3428.2007.03.024

[19] Zhang Y, Zhang P (2015). Machine training and parameter settings with social emotional optimization algorithm for support vector machine. Pattern Recognition Letters, 54, pp. 36-42. https://doi.org/10.1016/j.patrec.2014.11.011