

Research Article

Job Matching Analysis Based on Text Mining and Multicriteria Decision-Making

Fangning Liang ¹ and Xiangyu Wan ²

¹Faculty of Economics and Management, Langfang Normal University, Langfang 065000, Hebei, China

²Institute of Quantitative and Technological Economics (IQTE), Chinese Academy of Social Sciences (CASS), Beijing 100732, China

Correspondence should be addressed to Fangning Liang; 1301723@lfnu.edu.cn

Received 26 May 2022; Revised 26 August 2022; Accepted 27 August 2022; Published 21 September 2022

Academic Editor: Zaoli Yang

Copyright © 2022 Fangning Liang and Xiangyu Wan. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The postmanagement of an enterprise is to rationally and scientifically arrange the talents who meet the postrequirements and are suitable for the development of the enterprise in their suitable posts through testing, selection, and employment. The job recruitment information published on the recruitment website is the data that best reflects the market demand for data analysis talents. However, as online recruitment information is mostly presented in the form of text, this paper explores the information contained in the recruitment information through text mining technology, which has reference value for data analysis job seekers. Based on text mining and multicriteria decision-making, this paper studies job matching. The technical jobs under text mining are mainly distributed in the Internet industry, and the number of recruits accounts for 66.45%. Other industries have less demand for technical data analysis, all of which are below 15%. However, the market of business data analysis is scattered and extensive, and education still occupies the largest demand provider, accounting for 30.53%. However, real estate and media are closely followed, accounting for 24.28% and 22.34%, respectively. Besides these three industries, compared with technical talents, service industries will also need business data analysis talents more. In the stage of data collection, through the data analysis of the recruitment website, the job recruitment information is crawled, as well as other job information such as job salary, job qualifications, work experience requirements, and the number of recruits. This paper carries out the data collection stage and crawls the job recruitment information through the data analysis of the recruitment website. And, other position information, such as position salary, position qualification, and work experience requirements, as well as the number of recruits. The research of this paper can help the data released by the recruitment website to analyze the recruitment information and quickly reflect the market demand data for data analysis talents.

1. Introduction

The core of human resource management is the so-called job-matching. The human resource management strategy based on job-matching adopted by enterprises can realize the effective human resource management of enterprises to a certain extent. The traditional person position matching theory focuses on selecting the right person for a specific job, which has some defects. The postmanagement of the enterprise is to reasonably and scientifically arrange the talents who meet the postrequirements and are suitable for the development of the enterprise on their suitable posts

through testing, selection, and employment and carry out a series of management activities such as training, performance appraisal, salary feedback, and position adjustment, so as to realize the comprehensive management of the employees where the post is located [1, 2]. Person organization matching emphasizes the matching, mutual influence, and interdependence between employee characteristics and organizational characteristics, which is conducive to improving employee loyalty and organizational efficiency [3]. Enterprises should innovate the traditional mode of matching between people and positions, realize that the internal positions of enterprises are more inclined to the

matching between people and organizations, and maximize the effectiveness of enterprise human resource management. Organically combine the characteristics of the post with the characteristics of the individual to achieve the ideal effect of human resources management [4]. It includes four aspects: (a) each job has special requirements. (b) If an individual wants to be competent for a job, he must have certain knowledge, skills, and talents. (c) How to match the job characteristics with personal characteristics has a problem of suitability. (d) For everyone, job-matching means some kind of result.

With the improvement of enterprise's awareness of human resource management, the matching between people and posts has gradually become a new topic in the research of enterprise management. Through the matching between people and posts, the quality of enterprise's postmanagement can be comprehensively improved, and the smooth development of enterprise's management can be ensured. This is of great help to the improvement of enterprise's market competitiveness and the implementation of enterprise's strategic objectives. This is analyzed in depth below [5, 6]. The recruitment information of data analysis posts published on the recruitment website is the data that can best reflect the market demand for data analysis talents. However, as the online recruitment information is mostly displayed in the form of text, it is difficult for job seekers to search for the key information they want in the massive job information. Therefore, this paper explores the information contained in the recruitment information through text mining technology, which has reference value for data analysis job seekers [7]. Text mining technology involves many disciplines, including data mining, natural language processing, mathematical statistics, and visualization technology. Specifically, it refers to the process of using related technologies to discover and extract implicit knowledge from text data and form understandable information for users. It usually includes three stages: text preprocessing, text mining, and quality evaluation [8]. As an important carrier of information transmission in human society, the text shows a high degree of unstructured and discontinuous internal data relations. Each document takes words as the basic unit, and complex phrases and sentences are formed by the combination of words, thus forming the whole document. Not only are there ever-changing syntactic levels, but there are also many variants and synonyms, near meanings, ambiguities, and other relationships in words themselves [9, 10]. It is a necessary process of text mining to standardize and desensitize the original text data, and only keep the data useful for subsequent research.

Costa et al. discuss teacher evaluation from a comprehensive perspective of innovation. Based on the concept and method of multicriteria value measurement, it proposes a new teacher evaluation model. The model involves the whole academic activities and can be applied in different scientific fields and across different scientific fields. At the same time, respect its particularity [11]. Combined with the actual background of scientific research personnel performance evaluation, this paper establishes a performance evaluation index system. The performance of scientific researchers is

comprehensively evaluated by the combination of qualitative evaluation and quantitative evaluation. At the same time, the key information in the text is often hidden in many noisy data. Removing noise from text data is also the key to determine the quality of modeling results. In the data collection stage, by crawling the recruitment information of the data analysis post of the recruitment website, we can not only get the demand description text of each data analysis post. You can also get the location, industry, company size and attributes of the position corresponding to the enterprise that published the recruitment information, as well as other information about the position. Such as postsalary, posteducation, work experience requirements, and number of recruits. [12].

The research innovation of this paper is that it can help the data published by the recruitment website to analyze the recruitment information. Quickly reflect the market demand for data analysis talents. A multicriteria decision analysis model is used to comprehensively evaluate the performance of scientific researchers by combining qualitative evaluation and quantitative evaluation. This model can accurately get the performance evaluation value of researchers and then get the performance classification grade of researchers in the evaluation cycle according to the performance evaluation interval of researchers. Compared with the method of converting comments into scores, the algorithm of directly using performance comments for recommendation can make full use of the hidden information in comments. Describe users' multicriteria preferences and generate more satisfactory recommendation results.

2. Related Work

2.1. Research Status. In recent years, various comment based recommendation systems have emerged in academia. They transform the text comments given by users into valuable information for the recommendation process. According to the processing methods of user comments, these comment based recommendation algorithms can be divided into two categories: transformation scoring method and semantic vector method. At present, the technical entry threshold of data analysis posts is high. Moreover, the position of data analysis is not like that of business marketing. These posts need to be constantly recruited with fresh blood. For data analysis positions, employers are more inclined to recruit data analysis veterans with business experience or novices with better technology. And, the number of recruitment for data analysis posts is not large. Bryson et al. proposed that according to the different qualities and characteristics of individuals, the personnel should be arranged in the most suitable posts and the most suitable candidates should be selected for the posts, so as to "make the best use of people and things." There are two meanings: first, the postneeds its talents, that is, the postneeds employees with a certain quality and ability; second, people are suitable for their posts, that is, employees can be fully qualified for the post [13]. Mukherjee et al. proposed to explain the scope of responsibility of postwork and its position in the organization so that

employees can understand the main contents of work and carry out work within the scope of postresponsibility, so as to ensure the orderly work of each post [14]. Santis et al. proposed that scientific and effective rational use of talents is the key to human resource management. The key point of its core is that the characteristics of employees are consistent with the responsibility characteristics of the post, the needs of employees and the incentive behavior of employees are consistent with the remuneration provided by the post and grasp the mutual integration and matching of personal characteristics and enterprise organizational characteristics [15]. Lozej proposed that in order to enable employees to quickly enter the end of the post, systematic training on the skills and related knowledge used in the postwork is also required during the training, so as to deepen employees' understanding and mastery of work skills and knowledge and improve their work ability [16]. Florez puts forward that with the change of the external environment, the matching degree of the post will also change. If the enterprise does not adjust the postdesign with the development of society, the ability of employees far exceeds the requirements of the original post, and their career expectations far exceed the satisfaction brought by work. Therefore, employees may choose to change jobs [17]. Zhang et al. put forward that job matching is based on enterprise strategy, employee quality, enterprise environment, enterprise scale, technical content, enterprise development, and other factors and realizes job selection through a series of process control and supervision of job analysis, design, configuration, training, planning, evaluation, incentive and restraint, adjustment, and so on [18]. Ren et al. put forward that enterprises and individuals are a whole and have common interests. What enterprises provide is a career platform. If employees are suitable for any position, they will try their best to arrange them in suitable positions, which will be of great benefit to personal career development. For enterprises to maximize the role of talents, enterprises will also get corresponding returns, so as to achieve a real win-win situation [19]. Naz et al. put forward that if employees do not make progress, make steady progress and fail to meet the new requirements of their posts, they will be eliminated by enterprises; Employees' own efforts will also affect the change of job matching. If employees perform very well in their posts, then enterprises are likely to adjust their higher positions [20]. Yager et al. put forward that the matching between people and organizations emphasizes the matching, mutual influence, and interdependence between the characteristics of employees and the characteristics of organizations, which is conducive to enhancing the loyalty of employees and improving the efficiency of organizations. Enterprises should innovate the traditional mode of matching people with posts, realize that posts within enterprises are more inclined to match people with organizations, and maximize the effectiveness of enterprise human resources management [21]. Jd et al. put forward that every job post has its corresponding job requirement specification, and only qualified personnel are qualified to enter this post. After entering the post, employees will receive a series of

postmanagement processes such as training, performance appraisal, and career planning, so as to contribute their own strength to the post and create benefits for the enterprise [22].

2.2. Research Status of Job-Matching Based on Text Mining and Multicriteria Decision-Making. Text mining is a new research field in recent years. Mainly from a large number of unstructured text information to find potential possible data patterns, internal relations, laws, development trends, etc. The process of extracting effective, novel, useful, understandable, and valuable knowledge scattered in text files and using this knowledge to better organize information. Text mining has a wide range of research fields, mainly involving natural language processing, machine learning, data mining, information retrieval, and many other contents. Researchers in different fields have different application purposes for text mining. This paper takes the web of science database as the data source. At present, data is basically stored in the form of structured data, while text mining deals with unstructured data mainly recorded in natural language. Both data mining and text mining are the process of discovering potentially useful association rules and development trends in large data sets. Only the processing objects are different, text mining is still applicable to data mining related technologies and methods after data preprocessing. Therefore, text mining can be seen as the combination of data mining and text information processing. It can be considered that text mining is the expansion of data mining to support more free data formats, and it is the breakthrough and innovation of data mining.

In view of this, this paper combs a large number of relevant literature and studies job-matching based on text mining and multicriteria decision-making. This paper uses natural language processing, text mining technology, and multicriteria decision-making to analyze the recruitment information of data posts and mine the core requirements of data posts. Data mining engineers need to conduct in-depth analysis and insight into user advertising behavior through massive data. Refine and discover business rules, guide the construction of recommended model features, locate product related data problems and analyze and optimize them. In the collection and preprocessing stage of data analysis job recruitment information, this paper selects the nationwide data in Internet recruitment to analyze the job recruitment data and divides the data into two parts for processing. One part is the text information of job requirements, and the other part is other short text information of the job, such as company attribute, workplace, industry, salary, education, and work experience. The enterprise itself is a whole, and various positions of the enterprise are in the whole enterprise. Therefore, in practice, in order to ensure the cohesion of the whole enterprise, the position setting should be consistent with the overall organizational design of the enterprise. There are many research studies on post-talent demand. With the development of information technology, text mining technology is gradually mature, which makes the information

acquisition and mining of postdemand free from the time-consuming and labor-consuming of manual collection and labeling. The postdemand analysis is automated through network information crawling and text mining technology, which makes the research on post-talent demand develop greatly.

3. Algorithm and Model of Text Mining and Multicriteria Decision Making

In this era, a large amount of information on the Internet is presented in the form of web pages, and the content of web pages is mostly expressed in text. The traditional information retrieval technology has been difficult to adapt to massive text data processing. How to efficiently extract data that meets the needs from large-scale text resources involves text mining. Big data in reality is often expressed as unstructured, overlapping, and dynamically changing text data. By combining large-scale text data without domain constraints with a knowledge base with domain constraints. We can give full play to the advantages of large-scale text data to deal with the problem of transforming unstructured data into structured data. Assign a numerical weight to each word in the document space to indicate the role of each word in describing a document d_i . The weight of words in a document d_i can be regarded as the coordinates of d_i in the document space. Here, the document d_i is regarded as a point in the document space, and then d_i can be regarded as a vector from the origin of the document space to the point determined by d_i . For example, we can directly extract the product features mentioned most frequently by users in reviews and then use them as classification criteria after screening and summarizing. This method has universal expansibility and can be applied to various scenarios, but whether the criteria are reasonable or not needs manual verification. Its main idea is to input one text in turn to judge the matching degree between the current text and the existing cluster, that is, to calculate the similarity between the current text and the cluster center. If the current text matches an existing cluster, the current text will be classified into the cluster, otherwise, a new cluster will be created. Take the text mining clustering of postrequirements in the following figure as an example to explain the algorithm principle of text mining. The flow chart of the text mining algorithm is shown in Figure 1.

Measuring the effect of the text mining algorithm is mainly to verify whether the result of the model is consistent with the category of the original text. Since the category of each text data of the text is unknown, the effect can be verified by manually labeling the category label of each text or selecting according to the topic category when taking the text data. Firstly, we manually split part of the comment data into small sentences according to punctuation and then judge whether the sentence has a clear criterion description tendency according to semantics. Finally, we divide the human into the comment subsets of different criteria and give them corresponding criterion labels, so as to obtain the training set of a classification model. Firstly, the performance evaluation index system is constructed, and then the quantitative and qualitative calculation of performance is

carried out, the weight of the evaluation index, performance benchmark, and performance upper limit are determined. On this basis, the performance value of the evaluation object can be calculated according to the given formula, and then the performance evaluation result of the evaluation object can be obtained according to the determined performance classification interval. The schematic diagram of multicriteria decision analysis model is shown in Figure 2.

After the initial fuzzy decision matrix $X^{(k)}$ is normalized, the decision matrix obtained is called the normalized fuzzy decision matrix $R^{(k)}$

$$R^{(k)} = \begin{bmatrix} x_{11k} & \cdots & x_{1nk} \\ \vdots & \ddots & \vdots \\ x_{m1k} & \cdots & x_{mnk} \end{bmatrix}, k = 1, 2, \dots, i. \quad (1)$$

Normalized decision fuzzy numbers have the following characteristics.

- (1) Decision fuzzy number is a dimensionless index
- (2) The decision fuzzy numbers are positive and have: $r_{ij} \in [0, 1]$
- (3) Decision fuzzy numbers are benefit indicators, that is, for decision makers, the larger the value, the better

To determine the weight vector of attributes, this process is divided into two steps.

The first step is the group integration of the scheme.

In order to obtain group opinions, you can use

$$r_{ij} = \sum_{k=1}^L v_k (r_{ijk}). \quad (2)$$

The individual language fuzzy decision matrix $R^{(k)} = (r_{ijk})_{m \times n}$ ($k = 1, 2, \dots, L$) is integrated into the group language fuzzy decision matrix $R = (r_{ij})_{m \times n}$.

Based on the group language fuzzy decision matrix $R = (r_{ij})_{m \times n}$, the comprehensive attribute value of scheme A_i can be expressed as follows:

$$Z_i(w) = \sum_{j=1}^n w_j r_{ij}. \quad (3)$$

The second step is to determine the weight vector of the attribute.

If the weight information of the attribute is completely known, the schemes can be sorted according to the attribute value $Z_i(w)$ ($i = 2, 2, \dots, m$). The attribute weight information considered in this model is a completely unknown decision-making problem, and the attribute weight needs to be determined in advance.

Let $\tilde{M} = (a_1, b_1, c_1)$, $\tilde{N} = (a_2, b_2, c_2)$ any two triangular fuzzy numbers, then

$$d(\tilde{M}, \tilde{N}) = \sqrt{\frac{1}{3} [(a_1 - a_2)^2 + (b_1 - b_2)^2 + (c_1 + c_2)^2]}. \quad (4)$$

Is the distance between \tilde{M} and \tilde{N} . In particular, when $a_1 = b_1 = c_1$ and $a_2 = b_2 = c_2$, \tilde{M} and \tilde{N} degenerate into two real numbers, there is

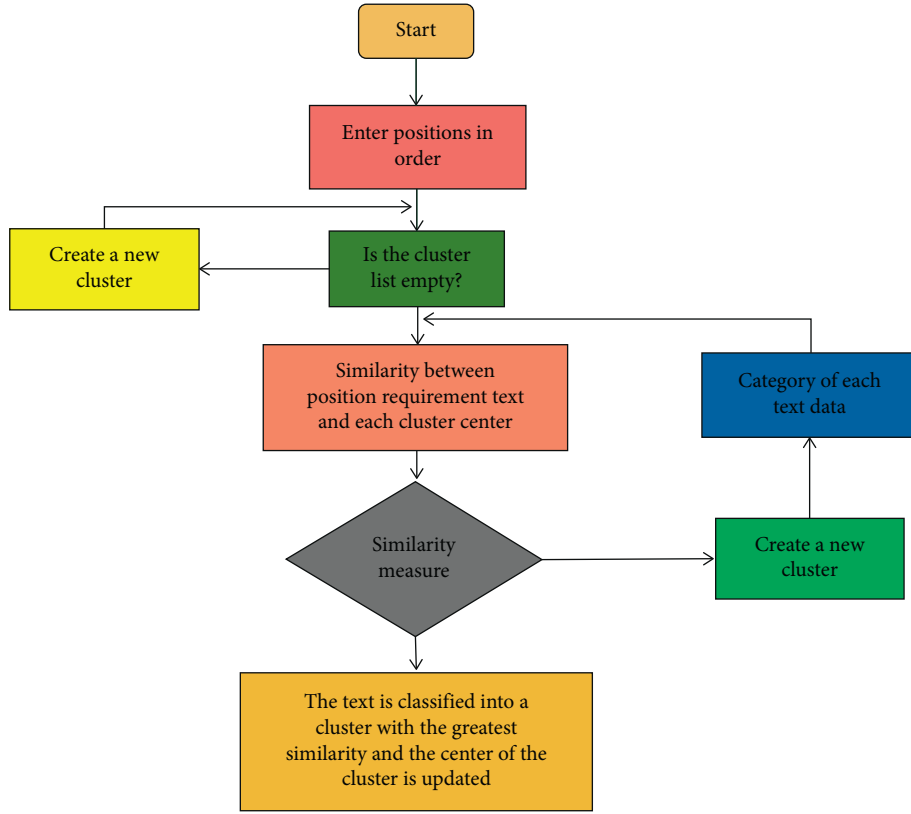


FIGURE 1: Flow chart of text mining algorithm.

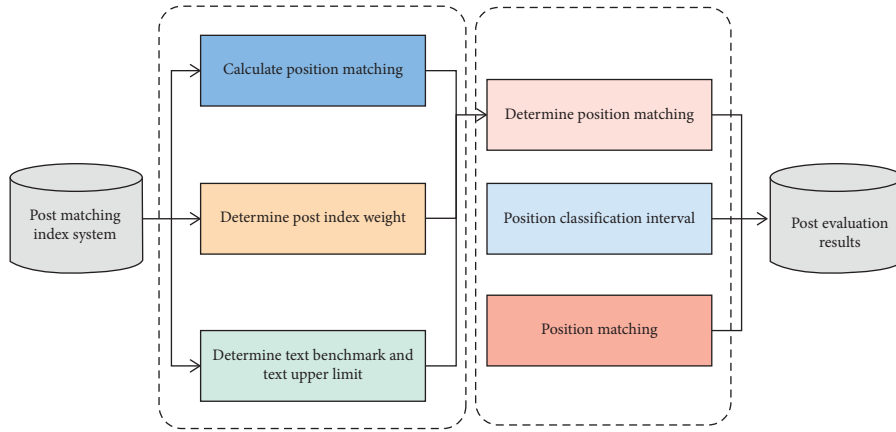


FIGURE 2: Schematic diagram of multi-criteria decision analysis model.

$$d(\tilde{M}, \tilde{N}) = \sqrt{\frac{1}{3} [(a_1 - a_2)^2 + (a_1 - a_2)^2]} = |a_1 - a_2|. \quad (5)$$

Based on the above theory, under attribute C_j , the deviation measure between scheme A_k and all other schemes is

$$d_{kj}(w) = \sum_{i=1}^m d(r_{ij}, r_{kj})w_j, i = 1, 2, \dots, m; j = 1, 2, \dots, n. \quad (6)$$

Then, for the j attribute C_j , the total deviation between all decision schemes and other decision schemes is

$$d_j(w) = \sum_{k=1}^m d_{kj}(w) = \sum_{k=1}^m \sum_{i=1}^m d(r_{ij}, r_{kj})w_j. \quad (7)$$

So as to construct the deviation function

$$d(w) = \sum_{j=1}^n \sum_{k=1}^m \sum_{i=1}^m d(r_{ij}, r_{kj})w_j. \quad (8)$$

Obviously, a reasonable weight vector $\omega = (\omega_1, \omega_2, \dots, \omega_n)^T$ should make the deviation $d(w)$ as

large as possible. Therefore, the following objective optimization models can be established.

$$\begin{aligned} \omega &= (\omega_1, \omega_2, \dots, \omega_n)^T, \\ \text{Max } d(\omega) &= \sum_{j=1}^n \sum_{k=1}^m \sum_{i=1}^m d(r_{ij}, r_{kj}) \omega_j. \end{aligned} \quad (9)$$

To solve the model, the following Lagrange functions can be constructed. Here, λ is Lagrange parameter.

Find the partial derivative of Lagrange function with respect to ω_j ($j = 1, 2, \dots, n$) and λ , and set $\partial L(\omega, \lambda) / \partial \omega_j = 0$ ($j = 1, 2, \dots, n$) and $\partial L(\omega, \lambda) / \partial \lambda$

Namely,

$$\begin{cases} \frac{\partial L(\omega, \lambda)}{\partial \omega_j} = \sum_{k=1}^m \sum_{i=1}^m d(r_{ij}, r_{kj}) + \lambda \omega_j = 0, \\ \frac{\partial L(\omega, \lambda)}{\partial \lambda} = \frac{1}{2} \left(\sum_{j=1}^n \omega_j^2 - 1 \right) = 0. \end{cases} \quad (10)$$

In fact, if a person thinks that one criterion is more important than other criteria when making a choice, he will often use stronger emotional words to express his satisfaction or dissatisfaction with this criterion after he has experienced the project. This situation means that when faced with a more valued criterion, people will usually judge whether the criterion meets their expectations by more strict standards, so it will be quite different from the average expectations of other users. Because there are not only dimensional differences between attribute indicators but also different scales, these differences will bring incommensurability, such as the most common attribute types, effective benefit attribute, and cost attribute. A benefit attribute refers to the attribute that the larger the attribute value is, the better. Cost attribute refers to the attribute that the smaller the attribute value is, the better.

4. Realization of Job-Matching

4.1. Overall Postsetting and Staffing. In the process of operation, the company adheres to the matching between personnel and posts and selects excellent talents for the company according to certain postrequirements under the mechanism of fair competition. Add fresh blood to the company's personnel team, fully mobilize the enthusiasm of personnel, give full play to the potential of talents, and better serve the development of the company. The human resources supervisor needs to have a certain basic ability to identify excellent talents, give full play to the enthusiasm of talents as much as possible according to the corresponding posts and personal judgment, widely collect opinions from all aspects when selecting talents, deeply understand the advantages and disadvantages of talents, make scientific and effective use of talents, fully understand talents as much as possible, select suitable talents according to specific posts, and give full play to the role of talents, and effectively serve the development of the company. When setting up posts, enterprises generally need to fully understand the text

mining and multicriteria decision-making posts that need to be set up, and fully understand responsibility requirements. They cannot set up posts simply because of events. After a comprehensive understanding of the postneeds, we should select talents scientifically and reasonably to maintain the scientific matching between personnel and posts. The post and employee quality of the enterprise are not invariable. With the change of social situation, the postdemand will change. With the enhancement or degradation of personnel quality, the postunder text mining will also be adjusted. Therefore, the enterprise needs to pay timely attention to the person job-matching and adjust the postsetting and personnel allocation in time, so as to ensure that the right person in the enterprise is in the right job. If we adopt the method of multicriteria decision-making and one post for life, it will hinder the development of the enterprise and reduce the work enthusiasm of employees. Therefore, we must make the person job-matching in a circular state.

4.2. Experimental Results and Analysis. This experiment takes the position of human resource manager in AS company as an example and calculates the matching degree of people and posts before adjustment. Single channel text clustering effect evaluation is to add category labels to each data through manual annotation and other methods and then use purity to evaluate clustering. The K-means clustering model is combined with the visual clustering model to obtain the best image clustering results. First, select the most important job requirements in the job description and give them weights according to the importance of each job requirement. The sum of the weights of each ability is 1. Then, all kinds of abilities of employees' quality are scored, with the range of 0–100 integers. The grades are divided into excellent 90–100, good 75–90, fair 55–75, poor 40–55, and poor 0–40. Calculate the average score scored by someone and fill it in the scoring column. Finally, calculate the product of each ability's weight and score and add it. The details are shown in Table 1. Whether the requirements of the position given by the enterprise match the abilities of the employees themselves. Often, employees' own ability determines their suitability for the post. The post requires employees to have the strong organizational ability, communication ability, talent training ability, and adaptability.

Select some of the most important factors in employees' career expectations and give them weights according to the importance of each factor. The sum of the weights of each ability is 1. Then, employees rate their job satisfaction, ranging from 0 to 100, and the grades are divided into excellent 90–100, good 75–90, fair 55–75, poor 40–55, and poor 0–40. Calculate the average score scored by others and fill it in the scoring column. Finally, the product of the weight of each occupation expectation and the score is calculated and added. The details are shown in Table 2.

Person-job-matching degree of other positions is calculated according to the above process. 1. Take the average for the situation of more than one post; 2. Because there are 110 specific positions of workers in 10 workshops of the

TABLE 1: Job requirements-matching degree of staff quality.

Postrequirements	Organizing ability	Communication ability	Talent cultivation ability	Strain capacity
Weight	0.5	0.4	0.3	0.2
Employee quality	Good	Good	Commonly	Commonly
Give a mark	84	82	66	62
Weight × scoring	42	32.8	19.8	12.4
Weighted result	77.6			

TABLE 2: Career expectation-job satisfaction matching degree.

Job expectancy	Salary situation	Lifting space	Team harmony	Working environment
Weight	0.3	0.4	0.3	0.2
Job satisfaction	Satisfied	Very satisfied	Common	Dissatisfied
Give a mark	82	85	62	47
Weight × scoring	24.6	34	18.6	9.4
Weighted result	74.5			

TABLE 3: Scores of person-job-matching of as company table a

Department	Post	Original person job-matching degree	New job-matching degree
Management	General manager	82	86
Department	General planner	71	78
	Director	74	81
Office	Correspondence clerk	66	73
Standardization	Manager	72	83
Department	Executive director	71	74

production planning department, the number is too large, a random survey method is adopted, and 6 people are selected from each workshop for investigation and calculation. Finally, the matching scores of people and posts before and after all positions adjustment in AS Company are obtained, as shown in Tables 3–6.

According to the literature review, it is found that the effect evaluation of text clustering using single pass is to add category labels to each data through manual annotation and other methods and then use purity to evaluate the clustering. The clustering objective function hopes to form a result of high similarity within clusters and low similarity between clusters. Take documents for example. Documents in the same cluster are more similar, while documents in different clusters are different. This kind of objective function is used as the internal standard of clustering quality. However, if the internal standard gets a good score, it does not mean that it can get good results in practical applications. The purity is calculated by the proportion of the coincidence part between the category label set generated by the clustering algorithm and the original category in the total sample size. The closer the purity is to 1, the more accurate the clustering result is. However, for the single pass clustering of postrequirements, we do not know which categories of data analysis posts are available in advance, so it is impossible to label them manually, so it is difficult to measure the effectiveness of the model with this method. Combine the K-means clustering model with the visual clustering model to obtain the best image clustering results. Firstly, select the best word vector model, input the results of the three word vector models into K-means, respectively, take the number of clusters K from 2

to 70, and conduct three experiments, respectively, for comparison. The experimental results are shown in Figures 3–5.

As can be seen from Figures 3–5, as the number of clusters increases from 2 to 70, the sum of squares of errors in the text clustering model combined with pretrained word2vec model and K-means algorithm is always the smallest of the three word vector models, which is significantly smaller than the model combined with job requirements word2vec model, word bag model, and TF-IDF. When the number of clusters K is small, the contour coefficient of using pretrained word2vec model combined with K-means clustering is significantly higher than that of other word vector models. When $k=2$, the contour coefficient of this model is 0.678, which is very close to 1, indicating that the sample clustering is more reasonable at this time. When $k=3$, the contour coefficient of pretrained word2vec model and K-means clustering model suddenly drops to 0.097. Therefore, the results of K-means clustering also verify the rationality of this post text being divided into two categories.

In this experiment, the industry distribution of jobs was investigated. The recruitment was concentrated in two areas: technical jobs and business jobs. The demand for data analysis technical jobs was evenly distributed in East China and South China, and the demand for technical jobs was more widely distributed in East China. Secondly, Central China, North China, and Southwest China are the second concentrated distribution areas of these two types of jobs. The difference is that the distribution of business jobs is Central China > Southwest China > North China, while the distribution of technical jobs is North China > Southwest

TABLE 4: Job-matching score of as company table b.

Department	Post	Original person job-matching degree	New job-matching degree
Finance	Cash register	65	83
Department	Bank settlement	61	65
Human resources	Manager	75	84
Department	Filer	72	78
Graduate school	Director of an institute	81	86
	Principal investigator	68	80

TABLE 5: Score of as company.

Department	Post	Original person job-matching degree	New job-matching degree
Chief engineer's office	Director	72	82
	Deputy director	72	77
	Manager	71	85
Production planning department	Workshop director	77	82

TABLE 6: Job-matching score of as company D.

Department	Post	Original person job-matching degree	New job-matching degree
Quality control department	Manager	73	85
	Administrators	71	75
Operating company	Chief of supply section	65	75
	Planner	72	72
Ministry of foreign trade	Foreign trade statistician	68	82
	Export salesman	73	77

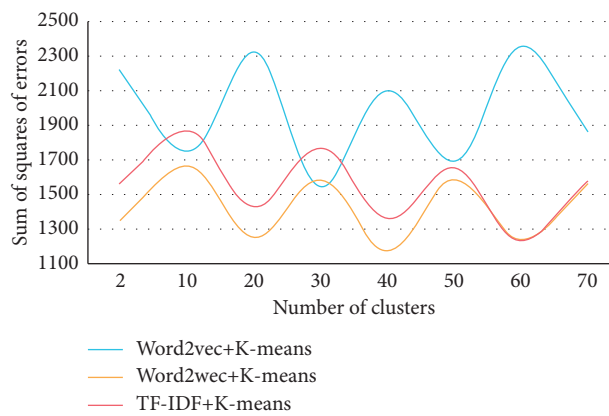


FIGURE 3: Different word vector models combined with K-means clustering.

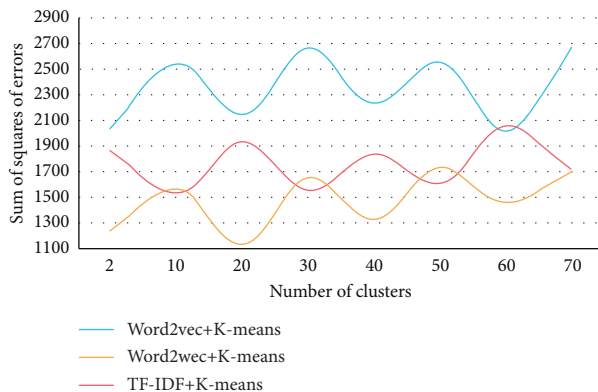


FIGURE 4: Different word vector models combined with K-means clustering.

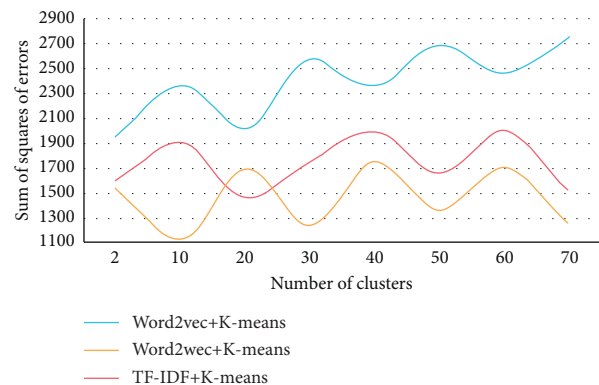


FIGURE 5: Different word vector models combined with K-means clustering.

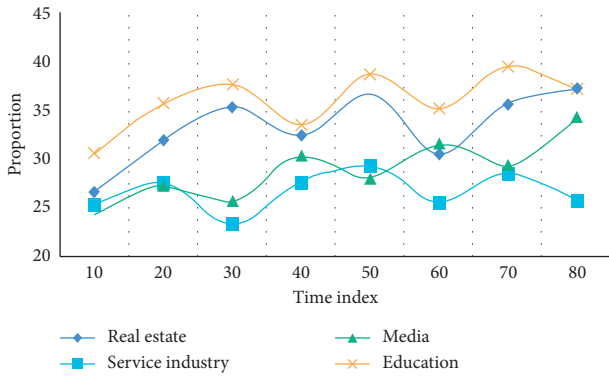


FIGURE 6: Industry distribution of Posts.

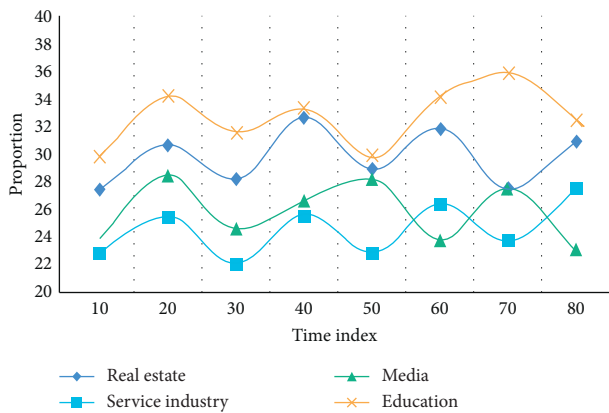


FIGURE 7: Industry distribution of Posts.

China > Central China. The less distributed areas are the northeast and northwest regions. In these two regions, there is more demand for recruitment in the northeast region, while the northwest region needs more technical data analysis talents. The experiments were compared twice, and the experimental results are shown in Figures 6 and 7.

From Figures 6–7, it can be seen that there are obvious differences in the industry distribution between technical posts and business posts. Technical posts are concentrated in the Internet industry, accounting for 66.45% of the total number of recruits. Other industries have less demand for technical data analysis, all below 15%; the market for business data analysis is relatively scattered and extensive. Education still accounts for the largest demand provider, accounting for 30.53%, but real estate and media follow closely, with 24.28% and 22.34% respectively. In addition to these three industries, compared with technical talents, the service industry will also need business data analysis talents more.

5. Conclusions

This paper studies the job-matching problem based on text mining and multicriteria decision-making. From the perspective of postresponsibilities, post-talents should have good comprehensive quality, have high requirements for “communication,” “coordination,” “language expression,” and other skills and have strong teamwork and decomposition ability. The multicriteria decision analysis model is

used to comprehensively evaluate the performance of scientific researchers by combining qualitative evaluation and quantitative evaluation. The model can get the accurate performance evaluation value of researchers and then get the performance classification grade of researchers in the evaluation cycle according to the performance evaluation interval of researchers. Text mining technology posts are mainly distributed in the Internet industry, accounting for 66.45%. Other industries have less demand for technical data analysis, which is less than 15%; the business data analysis market is relatively scattered and extensive. Education is still the largest demand provider, accounting for 30.53%, but real estate and media follow closely, accounting for 24.28% and 22.34%, respectively. In addition to these three industries, compared with technical talents, the service industry will also need more business data analysis talents. The post-management of the enterprise is to reasonably and scientifically arrange talents who meet the postrequirements and are suitable for the development of the enterprise to work in appropriate posts through assessment, selection, and use. The research of this paper can help the data released by the recruitment website to analyze the job recruitment information and quickly reflect the data of the market demand for data analysis talents. However, there are still some problems in this paper. The study did not reasonably design future employment goals, including whether to obtain more opportunities through improving education. How much is the difference between the ideal monthly salary and the actual monthly salary, and how much work experience is required before employment. Only on the basis of fully understanding the market situation can we better find jobs and choose jobs.

Data Availability

The data used to support the findings of this study are available from the corresponding author upon request.

Conflicts of Interest

The authors declare that there are no conflicts of interest.

References

- [1] M. Nützi, U. Schwegler, S. Staubli, R. Ziegler, and B. Trezzini, “Factors, assessments and interventions related to job matching in the vocational rehabilitation of persons with spinal cord injury,” *Work*, vol. 64, no. 1, pp. 117–134, 2019.
- [2] W. Lee and Y. J. Yoon, “Structural change in the job matching process in the United States, 1923–1932,” *European Review of Economic History*, vol. 26, no. 1, pp. 107–123, 2022.
- [3] V. Failla, F. Melillo, and T. Reichstein, “Entrepreneurship and employment stability - job matching, labour market value, and personal commitment,” *Journal of Business Venturing*, vol. 32, no. 2, pp. 162–177, 2017.
- [4] A. Bandara, “Youth labor market expectations and job matching in sub-Saharan Africa: evidence from school-to-work transition surveys,” *Applied Economics*, vol. 51, no. 8, pp. 762–780, 2019.
- [5] R. C. M. Beyer, “Job autonomy and education-skill matches of immigrant workers in Germany,” *Applied Economics Letters*, vol. 26, no. 16, pp. 1328–1332, 2019.

- [6] W. Lee, "Trade liberalization and the aggregate matching function in India," *Asian Economic Papers*, vol. 16, no. 1, pp. 120–137, 2017.
- [7] A. Adamecz-Vlgyi, P. Z. Lévy, and K. Brds, "Impact of a personalised active labour market programme for persons with disabilities," *Scandinavian Journal of Public Health*, vol. 46, no. 19_suppl, pp. 32–48, 2018.
- [8] S. Fujita, "Declining labor turnover and turbulence," *Journal of Monetary Economics*, vol. 99, no. NOV, pp. 1–19, 2018.
- [9] J. Devaro and O. Gürtler, "Advertising and labor market matching: a tour through the times," *Journal of Labor Economics*, vol. 36, no. 1, pp. 253–307, 2018.
- [10] H. Fu, R. Kleinberg, R. Lavi, and R. Smorodinsky, "Stability and auctions in labor markets with job security," *Economics Letters*, vol. 154, no. MAY, pp. 55–58, 2017.
- [11] C. Costa and M. D. Oliveira, "A multicriteria decision analysis model for faculty evaluation," *Omega*, vol. 40, no. 4, pp. 424–436, 2012.
- [12] D. Stijepic, "Trends and cycles in U.S. job mobility," *The Manchester School*, vol. 89, no. 2, pp. 203–222, 2021.
- [13] A. Bryson, J. Forth, and L. Stokes, "How much performance pay is there in the public sector and what are its effects?" *Human Resource Management Journal*, vol. 27, no. 4, pp. 581–597, 2017.
- [14] A. Mukherjee and L. Vasconcelos, "On the trade-off between efficiency in job assignment and turnover: the role of breakup fees," *Journal of Law, Economics, and Organization*, vol. 34, no. 2, pp. 230–271, 2018.
- [15] M. D. Santis, M. Florensa, and M. C. Gáname, "Job satisfaction of recent university graduates in economics sciences: the role of the match between formal education and job requirements," *Journal of Happiness Studies*, vol. 22, pp. 1–41, 2021.
- [16] M. Lozej, "Economic migration and business cycles in a small open economy with matching frictions," *Economic Modelling*, vol. 81, pp. 604–620, 2019.
- [17] L. A. Florez, "Job search inefficiency and optimal policies in the presence of an informal sector," *International Journal of Economic Theory*, vol. 15, no. 4, pp. 399–429, 2019.
- [18] Z. Zhang and W. Pedrycz, "Intuitionistic multiplicative group analytic hierarchy process and its use in multicriteria group decision-making," *IEEE Transactions on Cybernetics*, vol. 48, no. 7, pp. 1–13, 2018.
- [19] J. Ren and M. Lutzen, "Selection of sustainable alternative energy source for shipping: multi-criteria decision making under incomplete information," *Renewable and Sustainable Energy Reviews*, vol. 74, no. JUL, pp. 1003–1019, 2017.
- [20] M. N. Naz, M. I. Mushtaq, M. Naeem, M. Iqbal, M. W. Altaf, and M. Haneef, "Multicriteria decision making for resource management in renewable energy assisted microgrids," *Renewable and Sustainable Energy Reviews*, vol. 71, pp. 323–341, 2017.
- [21] R. R. Yager, "Categorization in multi-criteria decision making," *Information Sciences*, vol. 460, pp. 460–461, 2017.
- [22] A. Jd, B. Sw, and C. Smc, "Fuzzy best-worst method based on triangular fuzzy numbers for multi-criteria decision-making," *Information Sciences*, vol. 547, pp. 1080–1104, 2021.