A Method for Evaluating the Competitiveness of Human Resources in High-tech Enterprises Based on Self-organized Data Mining Algorithms

Sun Zhixin*
Anyang Vocational and Technical College, Anyang, Henan, 455000, China

***Abstract*—The level of human resources competitiveness of high-tech companies affects the efficiency and effectiveness of enterprises to a particular extent. To achieve sustainable development of high-tech enterprises, an evaluation method of human resource competitiveness of high-tech enterprises based on a self-organized data mining algorithm is proposed. The fuzzy clustering algorithm is used to select five first-level indexes for the evaluation of HR competitiveness of high-tech companies, including human capital power, human resources policy incentive power, and human resources performance manifestation power, and to construct the initial evaluation indicator setting. The self-organized data mining algorithm is used to identify the key attributes related to the human resource competitiveness of high-tech companies within the initial assessment indicator setup, reduce the complexity of the indexes and construct the final rating index system. The multi-level fuzzy evaluation method is applied to calculate the evaluation index weights and fuzzy evaluation matrix to obtain the assessment results of HR competitivity of high-tech enterprises. The experimental results show that the information contribution rate of the evaluation index system constructed by this method is higher than 95%, which can accurately evaluate the human resource competitiveness of high-tech enterprises.**

***Keywords*—Self-organized data mining algorithm; high-tech enterprises; human resources; competitiveness evaluation; multi-level fuzzy evaluation method**

I. INTRODUCTION

Human resources are the physical and mental workers who can promote social and economic development and create material and spiritual wealth for society [1, 2]. High-tech enterprises are organizations based on high technology. As creativity, information, and technology become more and more constituents of products, the knowledge content in products and services increases and the development of enterprises depends to a large extent on the ability to create knowledge and the level of technology possessed, thus human resources as carriers of knowledge and skills become strategic resources of enterprises [3-5]. The competitiveness of the human resources of a high-tech enterprise refers to the ability to integrate the strengths of the human resources of a high-tech enterprise with the process of human resource integration, which mainly emphasizes the ability to attract, invest, develop, utilize, and generate economic performance of HR of high-tech companies under the effective human resource management mechanism, to achieve the rapid and sustainable development of high-tech enterprises and open the gap with other high-tech zones [6]. However, high-tech enterprises still have the problem of a low overall level of human resource competitiveness in human resource management, so it is of considerable theoretical and practical importance to research how high-tech enterprises manage human resources and integrate human resource quality to bring into play the competitive advantage of enterprises.

Collins [7] examined how scholars in the field of strategic human resource management have utilized the resource-based view of the firm to contend that an HR strategy centered around high commitment can result in a competitive edge by establishing exclusive and valuable employee-centered resources at the firm level. Nevertheless, previous research has not clarified why variations in employee-centered resources remain among firms that adopt the same HR strategy, nor have they taken into account the significance of aligning employee-centered resources with other organizational abilities. The author suggested that the cognitive abilities of the CEO, social capital, and human capital can help clarify when a pursuit of a high commitment HR strategy leads to an increase in employee-centered resources and when firms can effectively handle and use them for competitive advantage. Utilizing Grey Relational Analysis (GRA) to analyze data from the CGSS and the China Health Statistics Yearbook for 2013 and 2015, Peng et al. [8] investigated the determinants of public satisfaction with the health system in China. The findings revealed that the percentage of total healthcare expenditure allocated by the government is the most critical factor for public satisfaction in both years. Moreover, out-of-pocket expenses were identified as a significant factor in 2013, while hospital beds per thousand population were crucial in 2015. The importance of healthcare workforce per thousand population increased from 2013 to 2015. The study's results indicated a transition in priority from economic affordability to more people-centric services in recent years. However, the research also highlighted persistent regional disparities and gaps that require attention in future healthcare reform endeavors. Geng and Hengxin [9] aimed to shed light on the evolution and contemporary environmental management practices of industrial parks in China. Industrial parks have played a pivotal role in China's economic advancement, particularly in terms of attracting foreign investment. The paper outlined the characteristics of industrial parks in China, delineated the principal site selection criteria, and highlights crucial factors for foreign investors. It also underscored the necessity of pursuing sustainable development in industrial parks, given their historical contribution to environmental pollution. The authors advocated for the adoption of eco-industrial development as a fresh developmental approach for industrial parks. Overall, the paper contributed to a more comprehensive comprehension of the current state of affairs and management outlooks concerning the growth of industrial parks in China.

In order to understand the competitiveness of enterprises, provide direction for human resource improvement in high-tech enterprises, and compare with competitors, the evaluation method of human resource competitiveness of high-tech enterprises based on a self-organized data mining algorithm is investigated. Follow the following approach for research:

1) Build an evaluation index system for human resource competitiveness of high-tech enterprises, comprehensively consider the weights and contributions of different indicators, and form a comprehensive evaluation result to provide a comprehensive and objective evaluation of human resource competitiveness.

2) Based on self-organizing data mining algorithms, key indicators are selected to determine human resource competitiveness evaluation indicators suitable for high-tech enterprises.

3) Evaluate the human resource competitiveness of high-tech enterprises by constructing a factor set, determining the weights of each indicator layer, constructing grade evaluation standards, determining membership relationships, constructing a fuzzy evaluation matrix, and determining a fuzzy subset vector;

4) Design experiments to verify that the proposed method can accurately evaluate the human resource competitiveness of high-tech enterprises.

The idea of self-organized data mining is the first milestone in developing mining theory and method. The self-organized data mining algorithm can use complete and incomplete induction algorithms to achieve the automatic selection of the optimal model. Through the research of this method, it is hoped to optimize human resource allocation, motivate employee motivation, improve employee satisfaction and retention rate, and promote the formation of innovation and learning culture. These benefits help enterprises improve their human resource management level, enhance competitiveness and sustainable development capabilities.

II. MATERIALS AND METHODS

A. *Construction of the Evaluation Index System for Human Resources Competitiveness of High-tech Enterprises*

In constructing the evaluation index system for the human resources competitiveness of high-tech enterprises, the fuzzy clustering algorithm is used to select evaluation indexes. Cluster analysis is a young branch of numerical taxonomy [10], which classifies indexes according to the degree of affinity of representative sample indexes in nature. In the assessment of the identifying the components of HR competitiveness of high-tech enterprises, because of the degree of reflection of various evaluation elements on the evaluation of HR competitiveness of high-tech enterprises, the boundaries of each other are not obvious. They have a certain degree of fuzziness [11], so employing fuzzy mathematical methods to deal with them is more appropriate.

The stages of using a fuzzy clustering algorithm to select the assessment indexes of HR competitivity of high-tech enterprises are as follows.

Step 1: Characterize the theoretical domain

The assessment elements of HR competitivity of high-tech enterprises are used to be classified as a thesis domain, $B=\left\{b_1,b_2,\cdots b_i,\cdots,b_m\right\}$, where b_i is characterized by a set of rating data, i.e., $b_i=\left\{b_{i1},b_{i2},\cdots,b_{ij},\cdots,b_{in}\right\}$.

Step 2: Determine the fuzzy relationship

The importance of various evaluation elements is fuzzily scored according to the specified scoring criteria to form the evaluation element scoring matrix $X\left(x_{ij}\right)_{m\times n}$. On this basis, the scoring matrix is normalized to calculate the correlation coefficient r_{ij} between different evaluation elements and establish the fuzzy similarity matrix R on the domain B. The most common method to determine the similarity coefficient r_{ij} is the closeness method. The closeness of u_i to u_j is $r_{ij}=N\left(x_i,x_j\right)$. When N takes the distance closeness, $r_{ij}=1-c\sum_{k=1}^{m}\left|x_{ik}-x_{jk}\right|$

Step 3: Calculate the fuzzy equivalence matrix

This fuzzy similarity relation matrix $R=\left(r_{ij}\right)_{m\times n}$, satisfies only self-reflexivity and symmetry but not transferability. To perform fuzzy clustering analysis, a Boolean multiplicative transfer closure operation is performed on R up to $R^k=R^{2k}\left(k=2,4\cdots.2^n\right)$. Taking $R^*=R^k=t\left(R\right)$, $t\left(R\right)$ is the fuzzy equivalence relation matrix, and the fuzzy equivalence relation matrix $t\left(R\right)$ satisfies self-reflexivity, symmetry, and transferability [12].

Step 4: Clustering

The fuzzy equivalence matrix $r\left(R\right)$ is used to find its intercept matrix under different thresholds ϕ, and the intercept matrix ϕ is used to analyze the similarity of the evaluation elements [13, 14] so that the indexes of each evaluation element can be clustered.

Step 5: Determine the optimal threshold ϕ

The intercept matrix of the fuzzy equivalence relation matrix $t(R)$ under different thresholds ϕ is $t(R)_\lambda = \lambda(r_{ij})$, where

$$\phi(r_{ij}) = \begin{cases} 1, r_{ij} \geq \phi \\ 0, r_{ij} < \phi \end{cases} \tag{1}$$

Different values of ϕ will get different truncation matrix R_ϕ. Generally, when ϕ gradually decreases from large to small, the classification becomes coarser from fine, forming a dynamic cluster analysis picture. Through this cluster analysis picture, the best ϕ value is selected according to the actual problem's needs, the classification objects' master-slave ranking can be realized, and the identification and analysis of the components of the HR competitiveness of high-tech companies can be realized.

Based on the principles of comparability, operability, the composition of qualitative and quantitative, dynamism, comprehensiveness, and complexity [15-18], the comprehensive index system of initial high-tech enterprises' HR competitiveness is constructed by combining the specifications of high-tech companies themselves. This index system mainly consists of five primary indexes, 16 secondary indexes, and 32 tertiary indexes, as presented in Table I.

The human resource competitiveness of high-tech enterprises mainly consists of five parts: human capital power, human resource environment attractiveness, human resource policy incentive power, human resource investment competitiveness, and human resource performance manifestation power, with the goal of reflecting both the existing competitiveness and the future competitive potential. Human resource environment attractiveness, human resource policy incentive power, and human resource investment competitiveness are the direct manifestation of human resource competitiveness of high-tech zone, and human capital power is the bridge between human resource environment attractiveness, human resource policy incentive power, and human resource investment competitiveness and human resource performance manifestation power.

B. *Selection of Key Indexes Based on Self-organized Data Mining Algorithm*

Since the initial human resource competitiveness evaluation index system of high-tech enterprises contains a large number of indexes, using all the indexes to evaluate the competitiveness of high-tech enterprises will consume a lot of time and space resources, so in order to improve the evaluation efficiency, it is necessary to analyze the indexes in the initial human resource competitiveness evaluation index system of high-tech enterprises, identify the key attributes related to the competitiveness of high-tech enterprises, and use them as the basis to design the assessment setup of high-tech companies' human resource competitiveness. In order to improve the evaluation efficiency, it is necessary to analyze the indexes in the initial evaluation index system of high-tech enterprises, identify the key attributes related to the competitiveness of high-tech enterprises, and design the evaluation system of human resources competitiveness of high-tech enterprises based on this.

In the process of identifying the key attributes related to the competitiveness of HR in high-tech companies, the objective system analysis (OSA) way in the self-organized info analysis algorithm is used to extract the key attributes [19]. The objective system analysis method in the self-organized data mining algorithm automates the following steps objectively.

1) Divide the observed sample data into training and detection sets.

2) Generate the system to be selected at each stage by different variables and growing complexity [20].

3) Estimate the unknown parameters on the training set for the parameters of high-tech enterprises' human resource competitiveness evaluation system.

4) Select some optimal human resource competitiveness evaluation systems for high-tech enterprises at each stage using data from the testing set.

5) When repeating steps (2) to (4), the complexity of the assessment setup of HR competitivity of high-tech enterprises grows gradually as the number of HR of high-tech companies increases. Otherwise, the final optimal evaluation system of human resources competitiveness of high-tech companies is selected.

According to the basic principles of the OSA algorithm, the key attributes related to the competitivity of HR of high-tech enterprises are identified in the following steps.

(1) Divide the initial data set of HR competitivity evaluation indexes of high-tech enterprises W into two groups with the same sample size γ and λ so that $W = \gamma \cup \lambda$. The data lengths of γ and λ are both N, and $P = \{1, 2, \cdots, m\}$, $Q = \{1, 2, \cdots, N\}$.

(2) Let $k = 1$. For the i-th variable (initial evaluation index data of high-tech enterprise's human resource competitiveness), parameter estimation is performed by the least squares method on the full set of samples [21]to obtain:

$$x_i = \gamma_0 + \lambda x_0 \quad i \in P \tag{2}$$

Then perform parameter estimation using least squares on the data sets γ and λ, respectively, to obtain:

$$x_i^\gamma = \gamma_0^\gamma + \lambda^\gamma x_0 = \gamma_0^\lambda + \lambda^\lambda x_0 \quad i \in P \tag{3}$$

Calculate the minimum deviation criterion value [22].

$$\eta_{1i} = \frac{\sum_{k=1}^{N} \left(\frac{x_i^\gamma(k) - x_i^\lambda(k)}{x_i(k)} \right)^2}{N} \quad i \in P$$

Let $\eta_1 = \min(\eta_{1i})$

Let $k = k + 1$ (4)

TABLE I. INITIAL EVALUATION INDEX SYSTEM OF HR COMPETITIVENESS OF HIGH-TECH COMPANIES

Target layer	First-level indexes	Secondary indexes	Tertiary indexes
Evaluation index system of HR competitiveness of high-tech companies	Human capital	Number of human resources	Total number of employees
			Proportion of R&D personnel
			Proportion of foreign students
		Human resource quality	Proportion of personnel with bachelor's degree
			Proportion of personnel with master's degree
			Proportion of personnel with doctor's degree
			Proportion of personnel with intermediate professional titles
			Proportion of personnel with senior professional titles
	Attractiveness of HR and environment	Perfection of human resources market	Human resource flow system
			Human resources entrepreneurship service satisfaction index
			Free flow index of human resources
		Economic improvement of HR	Per capita disposable income
			Contribution rate of high-tech industry
			Internationalization trend
		HR life comfort	Housing accessibility
			Transportation convenience
			Air environment index
			Cultural Leisure Index
		Intelligence intensity of HR	Number of colleges and universities per square kilometer
			Amount of scientific study institutions per square kilometer
			Amount of college students per square kilometer
	Human resources policy incentives	Attraction of HR policy	
		Incentive degree of human resources policy	
		Incentive degree of human resource flow policy	
		Perfection of human capital property rights system	
	Human resources investment competitiveness	Perfection of education and training system	Training input index
			Training efficiency index
		Health Care Security Index	Health insurance index for medical treatment
			Social security rate
		Investment in innovation	Funds for technical development activities
			Per capita scientific research funds
	Human resources performance demonstration	Scientific and technological achievements	Independent intellectual property rate
			technical income
		Economic performance	Per capita GNP
			Gross industrial output per capita
			Per capita profit and tax
		External benefit system of human resources	

Within the m independent variables, take any k different attribute variables $x_i, x_j, \cdots, x_r \left(i, j \cdots, r \in P\right)$, least squares estimate the parameters on the full set of samples to obtain a system of k -Formulas.

$$\begin{cases} x_i = \gamma_{01} + \gamma_{11} x_i + \cdots \gamma_{(k-1)1} x_r + \lambda_1 x_0 \\ x_j = \gamma_{02} + \gamma_{12} x_i + \cdots \gamma_{(k-1)2} x_r + \lambda_2 x_0 \\ \vdots \\ x_r = \gamma_{0k} + \gamma_{1k} x_i + \cdots \gamma_{(k-1)k} x_r + \lambda_k x_0 \end{cases} \tag{5}$$

Then carry out the parameter estimation for $x_i, x_j, \cdots, x_r \left(i, j \cdots, r \in P\right)$ using least squares on γ and λ, respectively, to obtain:

$$\begin{cases} x_i^{\gamma} = \gamma_{01}^{\gamma} + \gamma_{11}^{\gamma} x_i + \cdots \gamma_{(k-1)1}^{\gamma} x_r + \lambda_1^{\gamma} x_0 \\ x_j^{\gamma} = \gamma_{02}^{\gamma} + V_{12}^{\gamma} x_i + \cdots \gamma_{(k-1)2}^{\gamma} x_r + \lambda_2^{\gamma} x_0 \\ \vdots \\ x_r^{\gamma} = \gamma_{0k}^{\gamma} + \gamma_{1k}^{\gamma} x_i + \cdots \gamma_{(k-1)k}^{\gamma} x_r + \lambda_k^{\gamma} x_0 \end{cases} \tag{6}$$

$$\begin{cases} x_i^{\lambda} = \gamma_{01}^{\lambda} + \gamma_{11}^{\lambda} x_i + \cdots \gamma_{(k-1)1}^{\lambda} x_r + \lambda_1^{\lambda} x_0 \\ x_j^{\lambda} = \gamma_{02}^{\lambda} + \gamma_{12}^{\lambda} x_i + \cdots \gamma_{(k-1)2}^{\lambda} x_r + \lambda_2^{\lambda} x_0 \\ \vdots \\ x_r^{\lambda} = \gamma_{0k}^{\lambda} + \gamma_{1k}^{\lambda} x_i + \cdots \gamma_{(k-1)k}^{\lambda} x_r + \lambda_k^{\lambda} x_0 \end{cases} \tag{7}$$

Calculate the minimum deviation criterion value for each system of Formulas.

$$\eta_{ij\cdots r} = \frac{\eta_{ki} + \eta_{kj} + \cdots + \eta_{kr}}{k} \tag{8}$$

In Formula (8), calculate $\eta_{ki}, \eta_{kj}, \cdots, \eta_{kr}$ according to Formula (4), denoted as $\eta_k = \min\left(\eta_{ij\cdots r}\right)$.

(4) Compare η_k with η_{k-1}, if $\eta_k \leq \eta_{k-1}$, go back to step (3); otherwise, stop the algorithm and record the minimum deviation criterion value $best\eta = \eta_{k-1}$. The variables in the set of Formulas corresponding to the minimum deviation criterion value of η_{k-1} are the characteristic variables of the system [23]. The attributes corresponding to these characteristic variables are the key attributes for evaluating the competitiveness of HR of high-tech enterprises.

The obtained key attributes of the human resource competitiveness of high-tech companies are used to construct the final applied assessment index setup of HR competitiveness of high-tech enterprises, according to which the fuzzy comprehensive assessment process is applied to design the assessment project of human resource competitiveness of high-tech enterprises and obtain the evaluation results of human resource competitiveness of high-tech enterprises.

C. Evaluation of Human Resource Competitiveness of High-tech Enterprises

In the process of evaluating the competitiveness of human resources in high-tech enterprises, there are several elements to be assumed, and they are able to be parted into various layers and groups [24]. When conducting a comprehensive evaluation of such an object, in order to facilitate the distinction of the act and effect of the ultimate assessment of distinct elements and to receive the data that all elements collected more comprehensively, the number of elements is dividable to various groups based on particular features [25], with a small number of factors in one category and a comprehensive evaluation of each type first, followed by a high-level synthesis of the evaluation results between the classes. Therefore, the assessment of HR competitivity of high-tech enterprises is achieved by a multi-level fuzzy comprehensive evaluation method.

a) Construction of Factor Sets

The theoretical factor domain U is divided into m disjoint subsets according to different attributes, i.e., $B = \bigcup_{i=1}^{m} B_i$, and $B_i \cap B_j = \phi\left(i \neq j\right)$, and $B = \left\{B_i\left(i = 1, 2, \cdots, n\right)\right\}$ is called the first-level indicator set; each first-level indicator is further divided and $B_i = \left\{B_{ij}\left(i = 1, 2, \cdots, n; j = 1, \cdots, m\right)\right\}$ is called the second-level indicator set. Similarly, a set of indicators can be obtained at three or even more levels.

b) Determination of the Weights of Each Index Layer

The design of the weights of each level of indexes mainly uses the hierarchical analysis method to calculate the weights of tertiary, secondary, and primary indexes, in turn. The comparison matrices are constructed for different indicator levels, and the weights of each indicator are calculated [26], and then the consistency test is performed. If they pass the test, the weights can be determined; if they do not pass the test, the comparison matrix is reconstructed, the weights of each indicator are calculated, and the consistency test is conducted again until they pass the test so that the weights of each indicator can be scientifically determined.

c) Construction of Grade Evaluation Criteria

The evaluation level set is a collection of various evaluation results that evaluators may make about the competitivity of HR in high-tech enterprises, i.e., $V = \left\{v_1, v_2, \cdots, v_e\right\}$.

d) Determination of the Affiliation Relationship to Construct the Fuzzy Evaluation Matrix

The fuzzy assessment matrix for every element of the second level is created based on the third-level indexes. The factors of the third level index B_{ijk} are evaluated according to the rank of the evaluation set $V = \left\{v_1, v_2, \cdots, v_e\right\}$, and the affiliation degree of B_{ijk} to v_e is evaluated by $R_{ie}\left(i = 1, 2, \cdots, n; \mathrm{e} = 1, 2, \cdots, f\right)$, thus forming the evaluation matrix of the factors of the second level index R_{ij}.

$$R_{ij} = \begin{bmatrix} R_{11} & \cdots & R_{1e} \\ \vdots & \vdots & \vdots \\ R_{i1} & \cdots & R_{ie} \end{bmatrix} \tag{9}$$

R_{ij} denotes the fuzzy matrix corresponding to the factors of the 3rd level subset under the factor i at the 2nd level, i shows the number of elements of the 3rd level indexes under the 2nd level indexes, e is the number of human resource competitiveness level of high-tech enterprises, where: $R_{ie}\left(i = 1, 2, \cdots, n; \mathrm{e} = 1, 2, \cdots, f\right)$ denotes the affiliation degree of the f-th level comments made to the i-th evaluation index, i.e., later than using professionals to score the assessment outcomes, it can get v_{ie} level comments v_e for the i-th evaluation index, then:

$$R_{ie} = \frac{v_{ie}}{h} \tag{10}$$

Where, h is the number of experts.

e) Determination of the Fuzzy Subset Vector

The single-factor assessment matrix R_{ij} and the weight vector A_{ij} of the three-level index set $U_{ij}=\{U_{ijk}\}$ are used to do the fuzzy comprehensive assessment. The fuzzy integrated assessment vector of the three-level indexes is attained and normalized ("~" indicates normalization, same below).

$$P_{ij}=A_{ij}\cdot R_{ij}\sim\left(b_{ij}^{e}\right),\left(i=1,\cdots,n;\ j=1,\cdots,m;\mathrm{e}=1,2,\cdots,f\right) \tag{11}$$

Where, p_{ij}^{e} is the value of the affiliation function of the comprehensive evaluation P_{ij} taken at the e-th comment. The fuzzy operation "$\cdot$" here uses the $M\left(\cdot,\underset{+}{\wedge}\right)$ operator. That is, $p_{ij}^{e}=\overset{n}{\underset{+}{\wedge}}\left(a_{ij},r_{ij}^{e}\right),\left(i=1,\cdots,n;\ j=1,\cdots,m;\mathrm{e}=1,2,\cdots,f\right)$, where $\underset{a+b}{\wedge}=a+b-a\cdot b$.

In the same way, the comprehensive evaluation of the second-level indicator set and the first-level indicator set is done again. Eventually, the single-factor assessment matrix R and the weight vector A of the first-level indicator $B=\{B_i\}$ are used to do the fuzzy comprehensive assessment, and the comprehensive assessment vector P of the first-level indicator is attained and normalized; according to the above model, the fuzzy comprehensive evaluation results of human resources competitiveness of high-tech enterprises P_{ij}, P_i and P can be obtained, which are the fuzzy subset vectors on V.

$$\begin{cases}P_{ij}=A_{ij}\cdot R_{ij}=\left(P_{i1}\cdots P_{im}\right)\\ P_i=A_i\cdot R_i\\ P=A\cdot R\end{cases} \tag{12}$$

Thus, the results in assessing the HR competitiveness of high-tech enterprises are obtained.

III. RESULTS

A. Experimental Setup

To confirm the applied impact of the HR competitiveness evaluation method of high-tech companies according to the self-organized data mining algorithm studied in this article in the evaluation of the human resource competitivity of actual high-tech enterprises, eight high-tech enterprises in a first-tier city are used as the research objects. The overview of the research objects is shown in Table II.

The human resource competitiveness of each research subject is evaluated using the methodology of this paper, and the outcomes attained are shown hereby.

B. Info Clustering Results

In this paper, the fuzzy clustering algorithm is used to obtain the initial evaluation index system by clustering the research object's human resource-related data. Assuming that several human resource-related data points are random distribution, as shown in Fig. 1(a), where the noise points occupy 8% of the data, the number of clusters in this paper is set to 5. The clustering results of human resource-related data obtained by the method in the paper are shown in Fig. 1(b).

Analysis of Fig. 1 shows that the boundaries of various evaluation elements of the evaluation of HR competitiveness of high-tech enterprises are not obvious and have certain fuzziness. The method in this paper applies the principle of cluster analysis that "the indicators with similar properties are classified into one category, and the indicators with large differences are classified into different categories, so that the indicators within the category have high homogeneity, and the indicators between the categories have high heterogeneity," which can better shield the impact of noise and outlier data, and accurately reflect the spatial set characteristics of HR-related data. The initial assessment indicator arrangement of HR competitiveness of high-tech enterprises has been constructed, which has a positive impact on improving the application efficiency of human resource-related data.

C. Selection of Key Indexes

The method in this paper uses self-organized data mining algorithms to identify key attributes related to the competitiveness of high-tech enterprises on the basis of the initial evaluation index system of high-tech enterprise's human resource competitiveness, thus generating the final used evaluation index system for high-tech enterprise's human resource competitiveness, as shown in Table III.

In order to analyze the scientific of the evaluation index system of HR competitiveness of high-tech companies finally used, the principle of data variance describing the information content of evaluation indexes is used as the basis to set the analysis criteria for the scientific of evaluation index system construction. In and S are taken as the information contribution rate of the filtered evaluation indexes to the initial evaluation indexes and the covariance matrix of the evaluation index data, respectively, then.

$$In=\frac{trS_s}{trS_h} \tag{13}$$

In Formula (13), tr is the trace, and s and h are the number of filtered evaluation indexes and the number of initial evaluation indexes, respectively.

Formula (13) can describe the ratio between the overall variance of the screened evaluation indexes and the overall variance of the initial evaluation indexes so that the information on the initial indexes described by the screened evaluation indexes can be obtained. In general, if the screened indexes can reflect more than 90% of the initial indexes, it means that the evaluation index system constructed by the screened evaluation indexes is scientific.

The overall variance of the assessment indicators screened by way of this paper and the overall variance of the initial 32 evaluation indexes are brought into Formula (13) to obtain the results of the scientific analysis of the evaluation index system constructed by way of this article, as presented in Fig. 2.

Analyzing Fig. 2, we can get that after identifying the key attributes of the evaluation indexes in the initial evaluation index system, and the information contribution rate reaches more than 95% when the number of evaluation indexes reaches 17, which indicates that the assessment indicator setup constructed by the method of this paper has high scientific. Although the information contribution rate also increases under the condition that the number of evaluation indexes continues to increase, the complexity and redundancy of the evaluation index system also increase.

D. Evaluation Results

Taking enterprise, A as an example, the evaluation index weights are calculated using this paper's method, and the results are shown in Fig. 3.

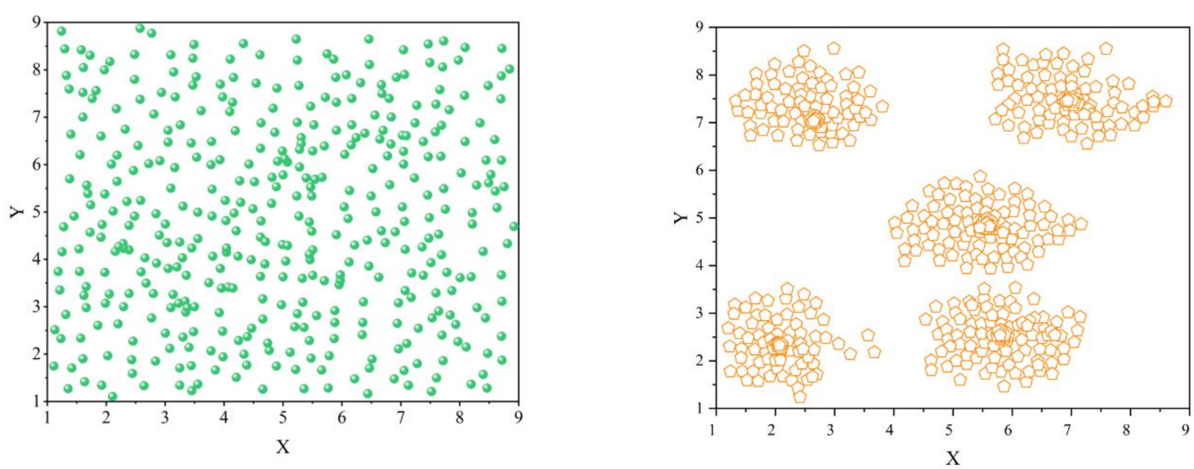
Analyzing Fig. 3, it can obtain that the weight of each level of the evaluation index is calculated using the method of this paper, and the highest weight of human capital power (B1) reaches 0.26. The lowest weight of human resource investment competitiveness reaches 0.15 (B4).

Fig. 4 shows the evaluation results of the methods in this article, studies [7], and [8] for each research object, as well as the comparison between the evaluation results of this article's method and the actual human resource competitiveness of each research object.

Analyzing Fig. 4, it can be seen that the human resource competitiveness of the selected research object does not have significant advantages and is generally at a moderate level. There is a significant deviation between the evaluation results obtained using the methods of studies [7] and [8] and the actual human resource development potential of the research object. This method can effectively obtain the evaluation results of the human resource competitiveness of the research object, and the evaluation results obtained are basically consistent with the actual human resource development potential of the research object. The major cause for the variance in the assessment outcomes of enterprise B maybe since the main direction of enterprise B is new materials and application technology and advanced manufacturing technology. The two technologies are related but relatively opposed to each other, so there is a certain duplication and redundancy in acquiring human resource information, which causes a certain deviation in the final evaluation results. Still, this deviation is within a manageable range. The above data fully demonstrate that the method of this paper can evaluate the competitivity of HR of high-tech enterprises more accurately.

TABLE II. OVERVIEW OF RESEARCH OBJECTS

Research object	Survey
Enterprise A	The main research direction is electronics and information technology
Enterprise B	Main research directions are new materials and application technology and advanced manufacturing technology
Enterprise C	The main research direction is modern agricultural technology
Enterprise D	The main research direction is aerospace technology
Enterprise E	The main research directions are new technologies for environmental protection and new energy and efficient energy-saving technologies
Enterprise F	The main research direction is nuclear application technology
Enterprise G	Main research directions are bioengineering and new medical technology
Enterprise H	The main research direction is marine engineering technology



(a) Random distribution of HR related data points.
(b) Clustering results.

Fig. 1. Clustering results of human resource-related data.

TABLE III. THE FINAL EVALUATION INDEX SYSTEM OF HUMAN RESOURCES COMPETITIVENESS OF HIGH-TECH ENTERPRISES

Target layer	First-level indexes	Secondary indexes	Tertiary indexes
Evaluation index system of HR competitiveness of high-tech companies	Human capital(B1)	Number of human resources(B11)	Total number of employees
			Proportion of R&D personnel
		Human resource quality(B12)	Proportion of personnel with master's degree
			Proportion of personnel with doctor's degree
			Proportion of personnel with senior professional titles
	Attractiveness of HR and environment(B2)	Perfection of human resources market(B21)	Human resource flow system
			Human resources entrepreneurship service satisfaction index
		Economic development of HR (B22)	Per capita disposable income
			Contribution rate of high-tech industry
	HR policy incentives(B3)	Attraction of HR policy(B31)	
		Incentive degree of human resources policy(B32)	
	Human resources investment competitiveness(B4)	Perfection of education and training system(B41)	Training input index
		Investment in innovation(B42)	Funds for technical development activities
			Per capita scientific research funds
	Human resources performance demonstration(B5)	Scientific and technological achievements(B51)	Independent intellectual property rate
			technical income
		Economic performance(B51)	Per capita GNP
			Per capita profit and tax

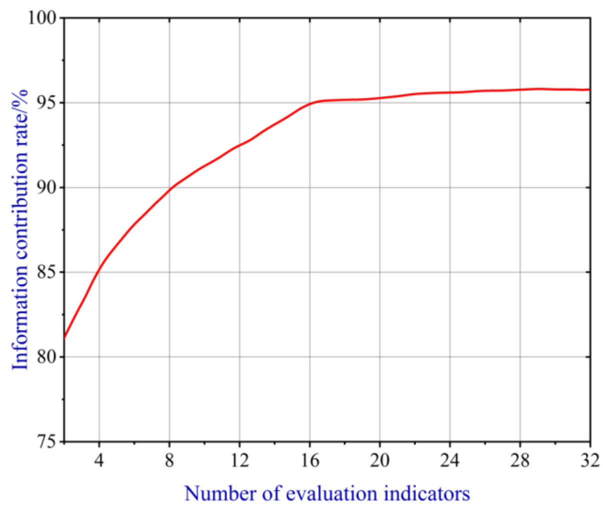


Fig. 2. Scientific analysis results of evaluation index system.

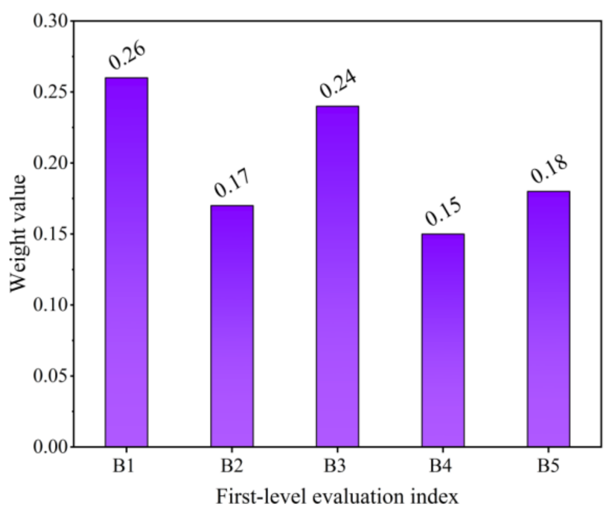
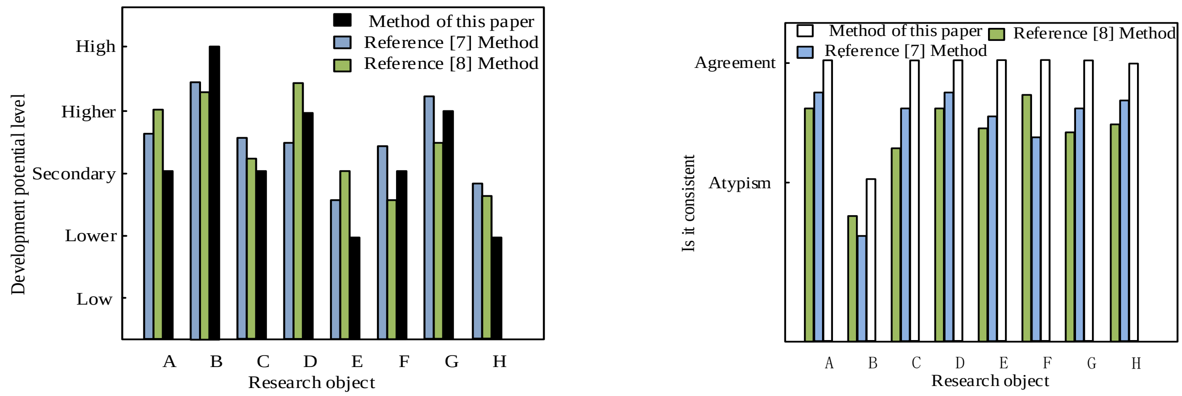


Fig. 3. Calculation result of weight of first-level evaluation index.



(a) Evaluation results of each research object

(b) Comparison with the actual situation.

Fig. 4. Evaluation results.

E. *Discuss*

In today's fiercely competitive business environment, the human resource competitiveness of high-tech enterprises is crucial for innovation and sustainable development. However, accurately assessing and measuring human resource competitiveness has always been a challenging task. Traditional evaluation methods often rely on subjective judgment and limited data, making it difficult to fully reflect the actual situation of the enterprise. Therefore, the evaluation method based on self-organizing data mining algorithm provides a more scientific and objective way for relevant research to evaluate the human resource competitiveness of high-tech enterprises.

The advantage of self-organizing data mining algorithm lies in its ability to automatically discover and select key evaluation indicators from a large amount of data, without the need for prior assumptions or manual intervention. This method can help identify factors that play a crucial role in human resource competitiveness, which may be overlooked or less noticeable in traditional methods. Through self-organizing data mining algorithms, key indicators that have a significant impact on the human resource competitiveness of high-tech enterprises can be extracted from massive data, providing a more accurate and comprehensive basis for evaluation.

In addition, evaluation methods based on self-organizing data mining algorithms can provide more reference and decision support. By analyzing and mining the selected key indicators, we can gain a deeper understanding of the correlation and degree of impact between these indicators. This in-depth analysis can help understand the complex relationships in human resource management and provide more targeted improvement and optimization suggestions for enterprises. For example, by identifying a positive correlation between employee satisfaction and innovation ability, companies can place greater emphasis on improving employee satisfaction to promote innovation and enhance competitiveness.

Overall, the evaluation method for human resource competitiveness of high-tech enterprises based on self-organizing data mining algorithms has obvious advantages and potential benefits. It can automatically discover and select key indicators from a large amount of data, provide comprehensive, objective, and accurate evaluations, and provide direction for improvement and optimization for enterprises. However, when applying this method, attention needs to be paid to issues such as data quality, algorithm selection, and comprehensive analysis of results. By overcoming these challenges and limitations, we can better evaluate the human resource competitiveness of high-tech enterprises, improve their competitiveness and sustainable development capabilities.

IV. CONCLUSION

This paper studies the assessment method of HR competitivity of high-tech enterprises based on a self-organized data mining algorithm that constructs an evaluation index system by using a fuzzy clustering algorithm and self-organized data mining algorithm and uses multi-level comprehensive fuzzy evaluation method to comprehensively evaluate the human resource competitiveness of high-tech enterprises, so as to determine the level of human resource competitiveness of high-tech enterprises, thereby judging the strength and weakness of human resource competitiveness of high-tech enterprises and discovering the defects of high-tech enterprises, which can provide a reference for measuring and improving the competitiveness of high-tech enterprises' human resources, and provide a reference for promoting the comprehensive, all-round and coordinated development of high-tech companies' HR competitivity. With the arrival of the era of Big data, more and more data exist in unstructured forms, such as text, image and audio data. Future research can explore how to use Natural language processing, image recognition, sound analysis and other technologies to incorporate these unstructured data into the evaluation model to obtain more comprehensive evaluation results.

REFERENCES

[1] M. Nyathi, R. Kekwaletswe, Realizing employee and organizational performance gains through electronic human resource management use in developing countries, African Journal of Economic and Management Studies, 14, pp. 121-134, 2023.

[2] Z. Xiaogang, C. Shuilin, L. Liqing, Research on the relationship among big data capacity, technological innovation and competitiveness of human resource service enterprises, Management Review, 33, pp. 81,2021.

[3] M.R. Azizi, R. Atlasi, A. Ziapour, J. Abbas, R. Naemi, Innovative human

resource management strategies during the COVID-19 pandemic: A systematic narrative review approach, Heliyon, 7, pp. e07233, 2021.

[4] B. Li, R.Y.M. Li, T. Wareewanich, Factors Influencing Large Real Estate Companies' Competitiveness: A Sustainable Development Perspective, Land, 10, pp. 1239, 2021.

[5] Z. Zhao, S. Yang, Y. Zhao, H. Chen, N. Dou, G. He, Z. Sun, Y. Yang, J. Luo, H. Gao, Status Quo and Equity Analysis of Human Resources for Health in China: Based on Five-year Data, health, 12, pp. 78, 2021.

[6] D.T. Nimfa, A.S.A. Latiff, S.A. Wahab, P. Etheraj, Effect of organisational culture on sustainable growth of SMEs: mediating role of innovation competitive advantage, Journal of International Business and Management, 4, pp. 01-19, 2021.

[7] C.J. Collins, Expanding the resource based view model of strategic human resource management, The International Journal of Human Resource Management, 32, pp. 331-358, 2021.

[8] X. Peng, X. Tang, Y. Chen, J. Zhang, Ranking the healthcare resource factors for public satisfaction with health system in China—based on the grey relational analysis models, International Journal of Environmental Research and Public Health, 18, pp. 995,2021.

[9] Y. Geng, Z. Hengxin, Industrial park management in the Chinese environment, Journal of Cleaner Production, 17, pp. 1289-1294,2009.

[10] W. Shunye, An improved k-means clustering algorithm based on dissimilarity, IEEE, pp. 2629-2633.

[11] X. Wang, J.H. Park, H. Yang, G. Zhao, S. Zhong, An improved fuzzy sampled-data control to stabilization of T–S fuzzy systems with state delays, IEEE Transactions on Cybernetics, 50, pp. 3125-3135,2019.

[12] T. Zhao, H. Cao, S. Dian, A self-organized method for a hierarchical fuzzy logic system based on a fuzzy autoencoder, IEEE Transactions on Fuzzy Systems, 30, pp. 5104-5115,2022.

[13] A. Mansouri, M.S. Bouhlel, Trust in Ad Hoc Networks: A New Model Based on Clustering Algorithm, Int. J. Netw. Secur., 21, pp. 483-493,2019.

[14] L. Mao, Q. Chen, J. Sun, Construction and optimization of fuzzy rule-based classifier with a swarm intelligent algorithm, Mathematical Problems in Engineering, 2020, pp. 1-12,2020.

[15] X. Cao, S. Guo, J. Lin, W. Zhang, M. Liao, Online tracking of ants based on deep association metrics: method, dataset and evaluation, Pattern Recognition, 103, pp. 107233, 2020.

[16] J. Fang, F.Y. Partovi, Criteria determination of analytic hierarchy process using a topic model, Expert Systems with Applications, 169, pp. 114306, 2021.

[17] F. Liu, M.-Y. Qiu, W.-G. Zhang, An uncertainty-induced axiomatic foundation of the analytic hierarchy process and its implication, Expert Systems with Applications, 183, pp. 115427, 2021.

[18] M. Mazzanti, Cultural heritage as multi-dimensional, multi-value and multi-attribute economic good: toward a new framework for economic analysis and valuation, The Journal of Socio-Economics, 31, pp. 529-558, 2002.

[19] W. Guangyao, W. Lizhen, Y. Peizhong, C. Hongmei, Minimal negative co-location patterns and effective mining algorithm, Journal of Frontiers of Computer Science & Technology, 15, pp. 366,2021.

[20] L.I. Ying, T. Yong, Research on the data mining algorithm based on association rules and similarity, 华南师范大学学报 (自然科学版), 53, pp. 121-127, 2021.

[21] J. Yuan, An anomaly data mining method for mass sensor networks using improved PSO algorithm based on spark parallel framework, Journal of Grid Computing, 18, pp. 251-261,2020.

[22] S. Ahmad, Adaptation of artificial intelligence in big data mining and its impact: a study, Solid State Technol, 63, pp. 2322, 2020.

[23] K. Yu, W. Shi, N. Santoro, Designing a streaming algorithm for outlier detection in data mining—An incremental approach, Sensors, 20, pp. 1261,2020.

[24] J. Yuguo, D. Asante, C. Dan, Z. Jie, Evaluation of low-carbon competitiveness based on a system evaluation method: A case study of three chinese steel companies, Mathematical Problems in Engineering, 2021, pp. 1-13,2021.

[25] L. Liu, D. Sun, Evaluation and Analysis of Comprehensive Competitiveness of Liaoning Province Based on Factor Analysis, 2020.

[26] C. Hongli, W. Yanyan, L. Xiuli, Z. Wenju, Competitiveness evaluation of chinese dairy industry based on accelerated genetic algorithm projection pursuit model, Mathematical Problems in Engineering, 2020, pp. 1-17,2020.