International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 17 (2025) pp. 1-17

https://doi.org/10.70917/ijcisim-2025-0281
Article

Research on the Innovation of Talent Cultivation Mode of Colleges and Universities Based on Big Data Technology Driven by School-Enterprise Cooperation

Lingli Zou *

School of Accountancy, Chongqing College of Finance and Economics, Chongqing, 402160, China; zoulingl2022@163.com

Abstract: Colleges and universities, as an important base for talent cultivation, are faced with the challenge of cultivating graduates adapted to the future needs of society and the market. The research firstly constructs the talent portrait of colleges and universities, carries out data collection and processing, for the problem of slow convergence speed of artificial bee colony algorithm for single peak problem, combines the improved artificial bee colony algorithm and fuzzy C-mean clustering algorithm to get the fuzzy C-mean clustering algorithm based on the improved artificial bee colony, and verifies it on the dataset, which finds that fuzzy C-mean clustering algorithm based on the improved artificial bee colony has more accuracy compared to traditional FCM with higher accuracy and the accuracy improvement reaches 90%. Through the improved clustering algorithm for cluster analysis, the students are finally classified into three types of students: self-disciplined learner type, night owl internet type and lack of planning type, on the basis of which a personalized teaching model for students' ability cultivation is designed. Finally, the BP neural network talent cultivation quality evaluation model was constructed and empirically analyzed. The study shows that the teaching mode can effectively improve the quality of talent cultivation in colleges and universities, and promote the innovative development of talent cultivation mode in colleges and universities and deepen the implementation of school-enterprise cooperation.

Keywords: behavioral portrait; fuzzy C-mean clustering; artificial bee colony; talent cultivation

1. Introduction

School-enterprise cooperation is a kind of cooperation mode created by schools and enterprises. Nowadays, the competition in the development of society is getting more and more intense, the education industry in order to better find the development, improve the quality of education and teaching, all the school-enterprise cooperation mode to make a brave attempt, and combined with the actual training of a large number of talents, the practicality of the talents and the comprehensive quality of the talents as the key point of personnel training [1-4].

Colleges and universities adopt school-enterprise cooperation talent training mode, on the one hand, can more accurately understand the market development profile, enrich their own educational resources, and cultivate students' innovative consciousness [5-6]. School-enterprise cooperation can promote the communication between colleges and universities and enterprises to form a consensus on talent training. Colleges and universities can make reasonable use of enterprise tools and facilities to organize and carry out practical teaching, explaining theoretical knowledge to students [7-8]; on the other hand, theoretical knowledge can also be applied to practice, so that students have a deeper understanding of the theoretical knowledge they have learned and deepen their memory [9-10]. In school-enterprise cooperation, schools need to establish a high-level teaching team to ensure that teachers can keep pace with the development of the times in the process of education and teaching, constantly update their cultural knowledge and concepts, and actively learn professional skills [11-13], or they can also invite professional workers to participate in the classroom teaching, share their work experience with the students, and teach the professional knowledge accumulated in the work for the students, so as to cultivate more big data talents needed by enterprises [14-16]. And with the continuous development of information technology, driven by school-enterprise cooperation, the innovative talent cultivation mode under the big data platform will further promote the cultivation of talents in colleges and universities through the adoption of big data analysis, intelligent technology and other means in teaching [17-20]. Under this combination, big data analytics information technology as a highly informative, mobile and interconnected new technology development pathway will be widely combined with the traditional education mechanism in colleges and universities, in-depth exploration of big data analytics information technology in the field of college and university education has great potential for development, and the innovative application of the college and university talent cultivation model has become a foreseeable inevitability [21-25].

As a new teaching mode, school-enterprise cooperation promotes the cultivation of talents in colleges and universities by combining the advantages of colleges and universities and enterprises, and at the same time meets the talent needs of enterprises. Literature [26] describes the significance of school-enterprise cooperation in cultivating talents, and analyzes the school-enterprise cooperation mode of cultivating applied talents by carrying out school-enterprise cooperative education and jointly revising talent cultivation programs, and points out the existing problems. Literature [27] emphasized that the innovation and entrepreneurship talent cultivation mode of school-enterprise cooperation can help transform the phenomenon of disconnecting teaching and practice in higher education and improve the comprehensive ability of college students in learning, innovation ability and practice. Literature [28] put forward suggestions related to the reform and practice of talent cultivation mode from the perspective of school-enterprise cooperation and industry-teaching integration in higher vocational colleges and universities. Literature [29] based on the concept of school-enterprise cooperation, through the investigation and analysis of the demand market of undergraduate talents, determined the objectives of talent cultivation, reformed and innovated the teaching mode, and formed the talent cultivation objectives with characteristics. Literature [30] discusses the talent cultivation mode of university-enterprise cooperative education in the context of new engineering disciplines, and puts forward a “talent cultivation mode” that runs through the whole process of talent cultivation objectives and talent cultivation programs. Literature [31] explains that the school-enterprise cooperation model will become an important strategic choice to promote the development of China's vocational education, which is conducive to the realization of a win-win situation for industry and learning, and puts forward an innovative school-enterprise cooperation operation mechanism to solve the problems faced in school-enterprise cooperation. Literature [32] analyzed the theory of school-enterprise cooperation in talent cultivation, and the results found that under the current conditions, the joint project development model is a kind of cooperation model acceptable to all parties, and affirmed the effectiveness of school-enterprise cooperation. Literature [33] introduced the talent cultivation mode of school-enterprise cooperation and its positive impacts and shortcomings, and based on the case study, it elaborated the value orientation, value realization and expectation achievement in the cooperative project, which provided reference value for the development of school-enterprise cooperation.

The traditional teacher-centered talent cultivation model has seriously hindered the development of teaching quality due to the lack of interactivity, and innovation as the core driving force of educational development, especially in the current era of big data, the traditional talent cultivation model has been unable to meet the needs of modern education, and the innovation of talent cultivation model of big data is imminent. Literature [34] examined teaching practices based on big data applications, aiming to enhance innovation in teaching, learning, sociality and technology, so as to provide students with innovative teaching and learning experiences. Literature [35] systematically introduces the science of big data and emphasizes the importance of big data based data analytics in enhancing innovation in teaching and learning paradigms, in addition to examining the advantages, benefits, limitations, and drawbacks of big data based analytics. Literature [36] discussed the significance and advantages of big data in college students' education and management, and explored the innovative applications of big data in terms of educational resource sharing, digital education and management. Literature [37] affirms the advantages that big data possesses in learning resources pushing, and analyzes the innovative application of big data technology in the network teaching mode of university courses in order to achieve the goal of expanding the function of network teaching mode with big data technology. Literature [38] emphasizes that big data technology can effectively solve the problems in English teaching and elaborates from the perspective of teachers and students, pointing out that big data provides effective teaching resources for teachers and students, and provides a broader perspective for English teaching. Literature [39] created an innovative teaching model for accounting majors based on big data technology, including innovative teaching objectives, teaching content, teaching methods and teaching evaluation, and outlined an innovative teaching model for accounting majors in the era of big data. Literature [40] reported the concept and characteristics of big data, pointed out the opportunities and challenges brought by big data for physical education teaching, and pointed out the development direction of university physical education in the era of big data. Literature [41] describes the contradiction between supply and demand and the basis of construction based on the problems of university-enterprise cooperation in talent cultivation in the context of big data, and discusses the advantages and disadvantages of several university-enterprise cooperation models.

In this paper, based on various campus data of colleges and universities, behavioral portrait of college students is carried out, and the artificial bee colony algorithm based on differential evolution is integrated into FCM, and the fuzzy C-mean clustering algorithm DEABC-FC based on improved artificial bee colony is proposed, and based on the algorithm, the effectiveness analysis is carried out. The DEABC-FC algorithm is used to develop the student clustering portrait analysis, on the basis of which the index system for evaluating the quality of student talent cultivation is constructed, and the BP neural network proposes the evaluation model of the quality of student talent cultivation based on the BP neural network and carries out the empirical analysis.

2. Design of university talent training model driven by school-enterprise cooperation

2.1. Construction of Talent Portrait in Colleges and Universities

2.1.1. Data collection and processing

Behavioral portrait of college students is to even out the whole picture of a student's information as comprehensively and meticulously as possible by mining the hidden information from the massive amount of college students' behavioral data. Behavioral profiling of college students involves various aspects of campus behavioral data. However, not all data are suitable for college students' behavioral portrait, selecting relevant behavioral information, eliminating unnecessary information, and determining the information data included in conducting the portrait.

Since most of the collected raw data are messy, incomplete and repetitive, the data need to be processed before the portrait is conducted, and the data processing of college student behavioral portrait includes data cleaning, data reconstruction, data integration and so on.

Data cleaning is to adjust or eliminate the data that do not meet the actual situation. Generally speaking, the collected data do not come from the same system, and there is inconsistency in the data format, so it is necessary to standardize the format of all the data and store them in the same database or data warehouse; in addition, there are often null values in the data stored in real databases, and it is common to use the substitution method, replacing the data with "0" or the average value of the item. The data reconstruction work, i.e. the generation of new data, is done. Reconstruction of data is the work of generating new fields and records. Data consolidation is the merging of related data.

2.1.2. Portrait label design

Labels are usually manually defined highly refined feature identifiers, and semanticization is an important feature presented by labels, which gives labels certain meanings so that people can easily understand these labels. Labeling is the core of college students' behavioral portraits, and labeled college students' behavioral portraits are convenient for both human understanding and computer processing.

For the time being, adding labels to college students is still based on manual collation, and the definition and representation of labels are generally determined by comprehensively collating the basic information of college students and various campus data. According to the campus behavioral data of college students, including students' performance data information, consumption data information, borrowing data information, access control data information and class check-in data information, etc., the labels of the portrait can be basically determined, including the basic attribute labels and dynamic attribute labels of college students.

2.1.3. Behavioral Profile Construction

When analyzing college student campus data, in order to analyze the whole picture of a student from a comprehensive perspective, it is necessary to conduct behavioral portraits of college students, so as to provide a data basis for analyzing students' academic performance, spending power and other information. Behavioral portrait is an image sketch of real students in campus activities, which puts a series of labels on students and realizes the labeling of students' campus behavioral information.

The core of college students' behavioral portrait modeling is to summarize a model that facilitates the analysis of students' information based on their basic information, performance information, consumption flow and other data, so that students' information can be labeled to facilitate the subsequent analysis of the students' situation, for example, "female, network engineering class, excellent grades, etc." can be used to describe the students' behavioral portrait. The behavior of college students is described according to the designed label attributes.

2.2. Improved artificial bee colony clustering algorithm for FCM student portraits

2.2.1. Improvement of the artificial bee colony algorithm

(1) FCM algorithm

The fuzzy C-mean clustering algorithm introduces a fuzzy affiliation matrix $\mathbf{U}$ on top of hard clustering, the matrix $\mathbf{U}$ contains n rows k columns, where the value of $u_{ij}\left(u_{ij}\in[0,1], i=1,2,\cdots,n; j=1,2,\cdots,k\right)$ values represent the extent to which a data object i belongs to a particular cluster j, i.e., fuzzy segmentation. Where $\mathbf{U}$ needs to satisfy Eq:

$$\sum_{j=1}^{k} u_{ij} = 1, i = 1,2,\cdots,n \tag{1}$$

Suppose the dataset we need to cluster is $\mathbf{X}=\left(\mathbf{x}_1,\mathbf{x}_2,\cdots,\mathbf{x}_n\right)$, where $\mathbf{x}_i$ is a D-dimensional data object, and we need to classify the dataset $\mathbf{X}$ into k classes. The clustering effect is generally measured by distance, and in this paper, Euclidean distance is used as the distance metric, then the objective function of FCM algorithm is:

$$J=\sum_{i=1}^{n}\sum_{j=1}^{k} u_{ij}^{m}\left\|\mathbf{x}_i-\mathbf{v}_j\right\| \tag{2}$$

where m is the weighted index and v_j is the corresponding j th division of the clustering center.

Using the Lagrange multiplier method, combined with the constraints (1), the derivation is obtained:

$$\mathbf{v}_j=\frac{\sum_{i=1}^{n} u_{ij}^{m}\mathbf{x}_i}{\sum_{i=1}^{n} u_{ij}^{m}} \tag{3}$$

$$u_{ij}=\frac{1}{\sum_{k=1}^{c}\left(\frac{\left\|\mathbf{x}_i-\mathbf{v}_j\right\|^2}{\left\|\mathbf{x}_i-\mathbf{v}_k\right\|^2}\right)^{\frac{2}{m-1}}} \tag{4}$$

The FCM algorithm first needs to determine the values of the number of clusters k, the fuzzy index m and the termination error ε. The clustering center v^t is randomly initialized according to the constraints, where t is the number of iterations and the current $t=0$. The next generation affiliation matrix $\mathbf{U}^{t+1}$, clustering centers $\mathbf{v}^{t+1}$, and objective function values $\mathbf{J}^{t+1}$ are obtained by iterating according to the above formula, if $\left\|\mathbf{J}^t-\mathbf{J}^{t+1}\right\|\leqslant\varepsilon$ then the iteration ends, otherwise it continues.

(2) Improvement of artificial bee colony algorithm

Artificial bee colony algorithm is a swarm intelligence algorithm that simulates the honey harvesting behavior of honey bees, according to the different division of labor, the bee colony is mainly divided into honey harvesting bees, observation bees, and detection bees [42]. A honey source corresponds to a honey harvesting bee, and the honey harvesting bee's job is to mine the honey source in anticipation of discovering a better source. The observation bee, on the other hand, speeds up the convergence of the algorithm by observing the information of the honey source brought back by the honey picking bee and selecting the honey source for mining with a certain probability based on the quality of the honey source. The remaining bees are scout bees, which are looking for available honey sources in the whole region to improve the global mining capability. The honey source $\mathbf{z}_i$ corresponds to be the solution of the objective function f, the quality of the honey source is measured by the fitness, and its value represents the goodness of the solution of the objective function, which is calculated as:

$$fitness_i = \frac{1}{1+f\left(\mathbf{z}_i\right)} \tag{5}$$

Through the study of artificial bee colony, it is found that the perturbation in the artificial bee colony algorithm is a kind of random and blind perturbation, the perturbation range is small and inefficient, and it does not utilize the information that the bee colony has already grasped, which leads to the slow convergence of the algorithm and is easy to fall into the local optimum. By changing the way of generating new nectar sources and adding the guidance of existing information, the optimization-seeking ability of the artificial bee colony algorithm can be strengthened.

The differential evolution DE algorithm is a heuristic stochastic search algorithm based on individual differences in the population, which outperforms other known stochastic search algorithms in terms of stability and convergence speed. Due to the good evolutionary ability of the differential evolution algorithm, the ideas of variation and crossover in differential evolution are introduced into the artificial bee colony algorithm to improve the convergence speed and accuracy of the artificial bee colony algorithm.

2.2.2. FCM algorithm based on improved artificial bee colony

FCM algorithm is widely used, simple and fast, but there are some shortcomings that restrict its development, such as the sensitivity to the initial point leads to unstable experimental results, large deviation, poor search ability, etc., which leads to clustering effect can not meet the clustering requirements [43]. The artificial bee colony algorithm has the advantages of insensitivity to the initial point, adaptability and robustness, and global search ability, while its shortcomings for the multi-peak problem, which is easy to fall into the local optimum, as well as the slow convergence speed at the later stage, have been well improved by introducing the ideas of variation and crossover in differential evolution. Therefore, we decided to combine the improved artificial bee colony algorithm and the FCM algorithm to obtain the DEABC-FC algorithm, and the advantages and disadvantages of the two algorithms can be well complemented to achieve better clustering results.

In the DEABC-FC algorithm, initialize the nectar $\mathbf{Z} = \left[\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_{SN}\right], \mathbf{z}_i = \left[z_{i1}, z_{i1}, \cdots, z_{im}\right]$, where $m = k*d$, i.e., each nectar $\mathbf{z}_i$ consists of cluster centers in $k*d$ dimensions, representing a fuzzy clustering division. In this case, the goal of the artificial bee colony algorithm is to find the honey source that makes the best clustering effect, so the objective function f of the bee colony algorithm is the objective function of clustering.

2.2.3. Algorithm effectiveness analysis

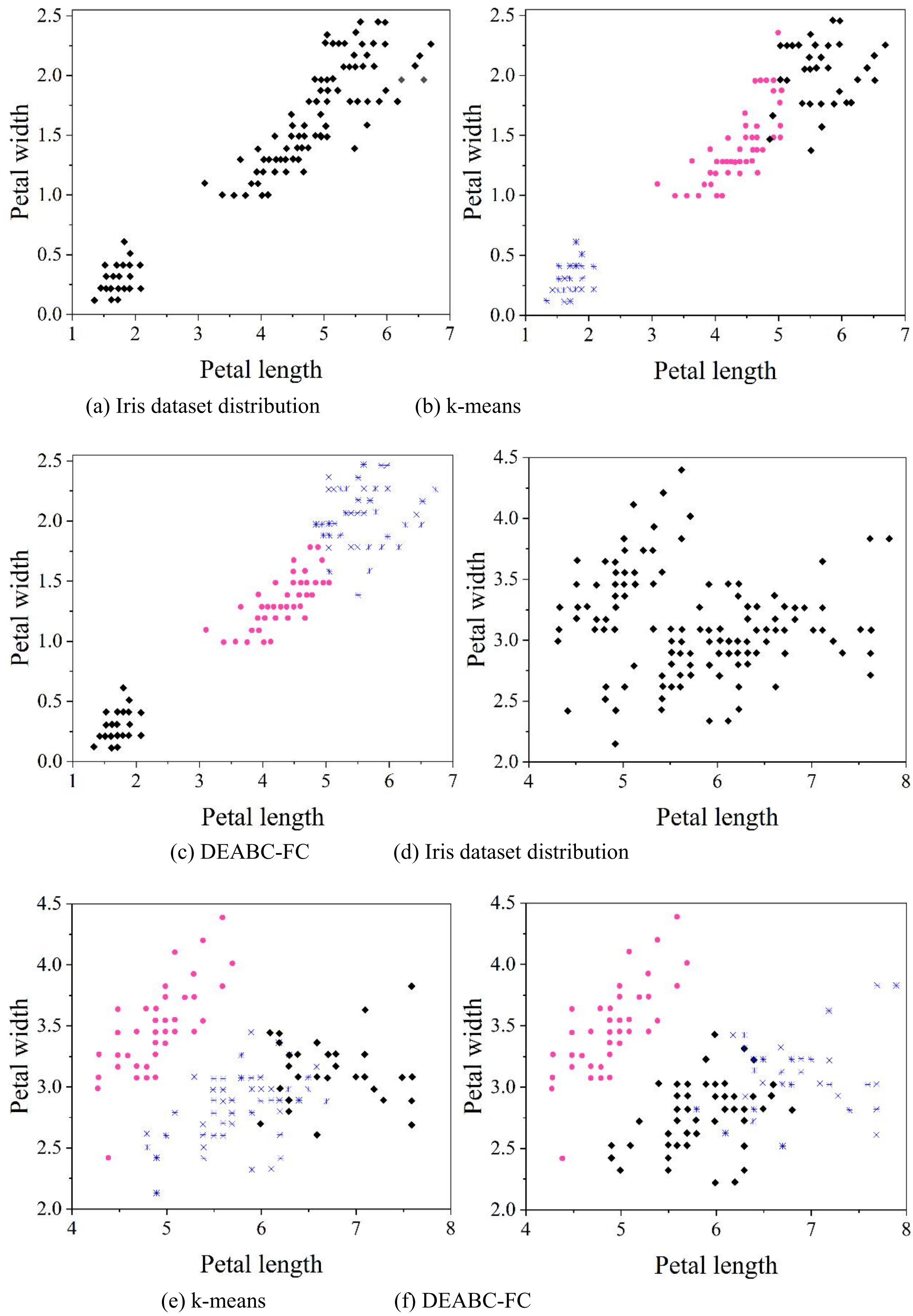
In order to verify the clustering effect of DEABC-FC algorithm, K-means, standard FCM algorithm, and ABCFCM are selected as the comparison algorithms in this paper, and 4 sets of datasets from UCI machine learning library are used for 20 independent tests. The attribute information of the datasets is shown in Table 1.

Table 1. Dataset properties.

Data set	Instance count	Data dimensions	Number of clusters
Iris	160	5	4
Wine	185	17	4
Glass	220	12	9
Wine Quality	4937	14	12

In order to show the clustering results of the algorithm, the attributes of the Iris dataset are selected as features for the experiment, and the clustering effect of the algorithm is shown in Fig. 1.The Iris dataset is an iris category dataset, which defines four attributes: calyx length, calyx width, petal length, and petal width. As can be seen from Figure 1, the clustering algorithm has no significant difference in the clustering effect for the cluster classes with obvious separation, but for the cluster classes with less obvious separation, the clustering results of the K-means algorithm will have a large gap due to the selection of different initial clustering centers, which yields poor results indicating that it falls into the local optimum and has a general ability to search globally. Compared with the hard clustering algorithm, the fuzzy clustering algorithm has improved the accuracy of the algorithm to a certain extent because the

DEABC-FC algorithm determines the clustering relationship by the degree of affiliation, and the clusters are not distinct from each other.



(a) Iris dataset distribution (b) k-means

(c) DEABC-FC (d) Iris dataset distribution

(e) k-means (f) DEABC-FC

Figure 1. Cluster rendering.

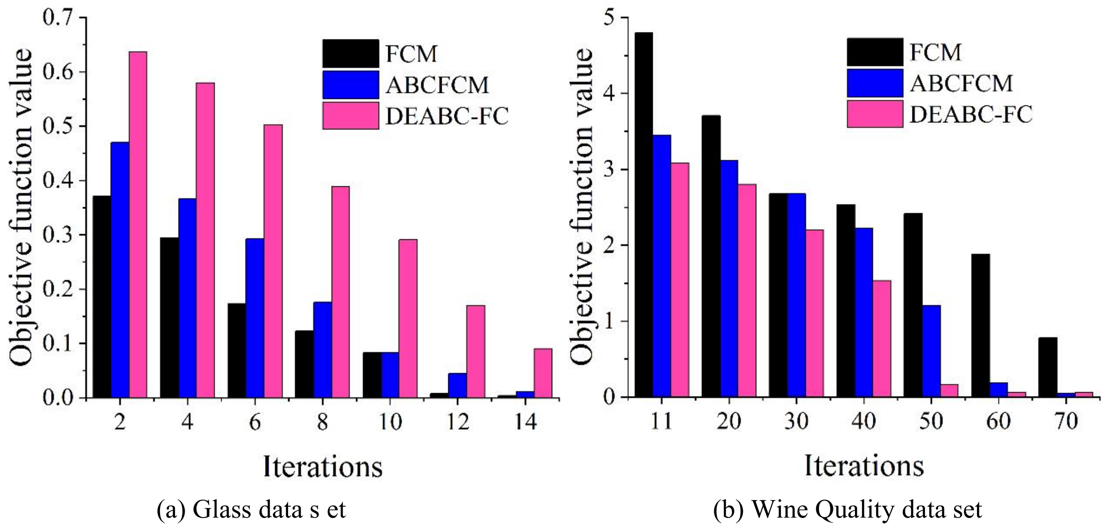
The results of the experimental comparison of different clustering algorithms are shown in Table 2. Analyzing the data in Table 2, it can be seen that on the three datasets, the differences in the algorithm models lead to different degrees of influence of the algorithms on the clustering accuracy. Among them, tested on the small number of Glass dataset, the K-means algorithm has the highest accuracy of 0.8425.

Since K-means is a clustering algorithm based on Euclidean distance, it has a greater advantage in the small number of datasets. With the increasing number of test datasets, by introducing the artificial bee colony algorithm, the ABCFCM algorithm has increased its search ability compared with the FCM algorithm, but the artificial bee colony algorithm has the disadvantage of being easy to fall into the local optimum, which leads to an insignificant increase in the accuracy of the ABCFCM algorithm. Neighborhood perturbation and contextual multi-armed gambling machine ideas are introduced to address the shortcomings of ABCFCM, and from the experimental results, it can be seen that the accuracy rate of DEABC-FC is significantly improved to reach 90% accuracy, and the improved algorithm can achieve better clustering results in datasets with a large number of instances.

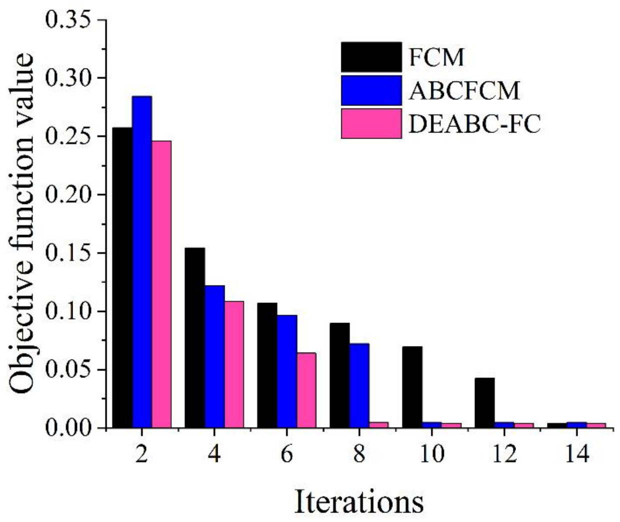
Table 2. Comparison of different clustering algorithms.

Algorithm	Data set	Precision
K-means	Wine	0.8334
	Glass	0.8425
	Wine Quality	0.7296
FCM	Wine	0.7852
	Glass	0.8054
	Wine Quality	0.8897
ABCFCM	Wine	0.8268
	Glass	0.7998
	Wine Quality	0.8953
DEABC-FC	Wine	0.8956
	Glass	0.8358
	Wine Quality	0.9123

The trend of ACU metrics for each algorithm on the dataset is shown in Figure 2. As can be seen from the figure, the lower the number of iterations, the stronger the convergence of the algorithm, and compared with the traditional FCM algorithm, the ABCFCM algorithm obviously has better convergence. While the DEABC-FC algorithm is tested on three datasets, on the Wine dataset and Glass dataset, the curve decreases fast with the number of iterations, and on the large number of datasets, the solution speed is optimal, and the number of iterations is the lowest, which indicates that it still has a good computational performance when dealing with high-dimensional data.



(a) Glass data s et (b) Wine Quality data set



(c) Wine data set

Figure 2. The curve of the objective function value versus the number of iterations on the dataset.

2.3. Student Cluster Portrait Analysis

In this study, based on the correlation between each feature value and academic performance, the hierarchical analysis method (AHP) was used to analyze the importance of features and determine new weights on the basis of ensuring the diversity of features, and to increase the differences between features on the basis of retaining the features as full as possible. Through the DEABC-FC clustering algorithm for centroid selection of multiple algorithmic experiments, it was finally found that when the clustering algorithm in accordance with the K = 3 clustering, to get the most significant differences in the characteristics of the student clusters, Cluster_0, 1, 2 a total of three clusters and centers as shown in Table 3, three-dimensional clustering as shown in Figure 3.

Night owls online students: corresponding to the cluster Cluster_0, such students are mostly male, compared to such students who have the best performance in three subjects when they enter the school, although they also read a certain amount of extracurricular books, but due to the lack of self-discipline and initiative, they are not willing to participate in public welfare activities, and they spend a long time on the Internet, especially at night when they are online the longest, and they have an irregular work and rest schedule, which leads to a regression of their academic performance after enrollment, reflecting the impact of the Internet on the students. This reflects that the negative influence of the Internet on students who lack self-control is even greater. This reflects that the Internet has a greater negative impact on students with a lack of self-control. If the guidance and supervision of students are strengthened to help them establish self-discipline in their study and living habits, it will have a greater impact on this group of students.

Lack of planning students: Corresponding to Cluster_2, the proportion of female students in this category is still relatively high, from the point of view of enrollment scores, this category of students has the lowest GPA, at the same time, after entering the school, they do not actively read extracurricular books or participate in extracurricular activities, they do not have high academic requirements, they are moderate in their thinking, lack of thinking about life planning, addicted to the Internet, and their academic scores are the lowest among the three groups. The academic performance is also the lowest among the three groups. Therefore, this kind of students can be listed as a key group of students to pay attention to the phenomenon of polarization of female students, can set up one-to-one support groups, through the attention, guidance and supervision of teachers, to help this kind of students to do a good job of life planning, set up learning goals, combined with the leadership and help of other peers, to develop good habits of study and life, and build up a certain degree of self-confidence in learning.

Table 3. Compare student profile clustering results.

Cluster Center	Cluster_0	Cluster_1	Cluster_2
Sex	Male	Female	Female
GPA for admission Chinese	3.609	3.4050	3.0630
GPA for admission Mathematics	3.4257	3.1550	2.8841
Admission GPA English	3.4924	3.2973	3.3841

Book borrowings	28.6157	38.4117	8.9574
Number of public service activities	5.3557	22.0217	7.1974
Activity credits	6.7357	19.1517	8.9074
Internet time/week	56.1358	27.5748	43.7374
Internet usage from 0 to 6 hours per week	11.2770	4.2717	7.2323
Internet usage from 7 to 12 hours per week	6.3531	3.1930	4.9294
Internet usage from 13 to 18 hours per week	17.5571	8.0127	12.1591
Internet usage from 19 to 24 hours per week	21.0257	12.1625	19.4688
Borrowed book type: Literature	0.4605	1.3695	0.1939
Book type: Economy	1.2214	0.5869	0.3115
Borrowed book types: Mathematics and Science, Chemistry	0.3953	0.3158	0.0854
Borrowed book type: lang	0.2647	1.0217	0.8434
GPA	2.7492	3.3521	2.6594

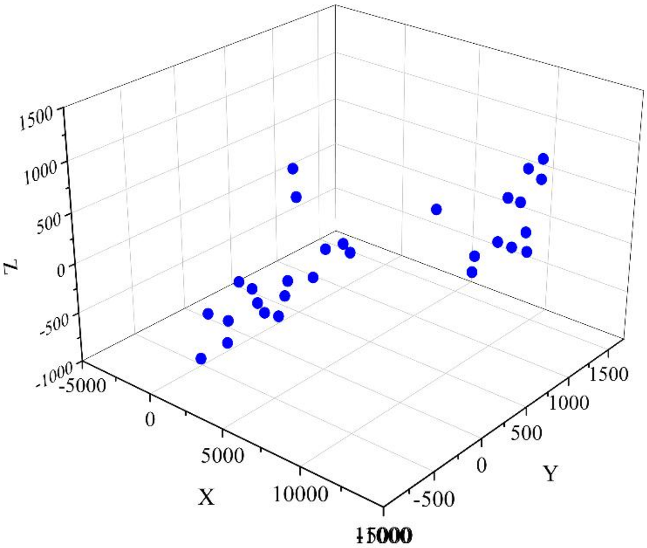


Figure 3. 3D Cluster Display.

2.4. Individualized teaching model design for capacity building of university students

Personalized analysis and portrayal of innovative behavior is conducive to the personalized cultivation of students' innovative ability, which is beneficial to the realization of the teacher's tailor-made teaching, and lays a solid foundation for the realization of personalized teaching. Therefore, under the guidance of the above ideas, the author designs a personalized teaching model for the cultivation of innovation ability of college students majoring in computer science and technology. The model consists of four stages: personalized teaching preparation, personalized teaching design and implementation, diversified teaching evaluation, and teaching reflection.

(1) Preparation for personalized teaching

In the pre-preparation, personalized teaching is divided into five parts, including learner analysis, teaching goal analysis, teaching content analysis, learning group distribution design and teaching tool design and implementation, here, the author analyzes and elaborates on these five parts from the perspectives of students, teachers and technology. For students, before the start of a course, it is necessary to initially diagnose the status of students' individual innovation ability, i.e., students initially diagnose their current level of innovation ability through a questionnaire survey based on the personalized education resource recommendation system of the students' innovation ability portrait and carve out a portrait of students' individual innovation ability.

(2) Personalized teaching design and implementation

The design of this teaching mode includes six teacher sessions: reviewing and reviewing, clarifying learning objectives and learning content, creating a context, personalized independent & cooperative learning process of individual-intra-group-inter-group knowledge map construction, complementary learning process of teacher-student complementary knowledge map construction, solving and answering the teaching content questions and practicing exercises. In the classroom, firstly, the teacher guides the review to review the old knowledge of the previous class, and secondly, based on the overall status of students' innovation ability in the class as understood by the recommender system, he or she specifies the learning objectives and learning content of this class; then a learning situation is introduced to stimulate students' curiosity and subjective initiative; then each student learns the new knowledge according to the educational resources recommended by the online recommender system to each of them, and after that, students will adopt the learned knowledge to After that, the students will digest and absorb the knowledge by constructing personalized independent & cooperative learning mode of individual-intra-group- inter-group knowledge mapping, and the learning knowledge also cultivates the students' abilities of independent learning, teamwork, communication, exploration and learning.

(3) Diversified Teaching Evaluation

Diversified teaching evaluation is an important link in realizing the education of innovative ability cultivation, and the author adopts the diversified evaluation method combining diagnostic-formative-summative evaluation. Diagnostic evaluation is adopted in the preliminary preparation for personalized teaching to realize the initial diagnosis of innovative ability. Then after a period of personalized teaching oriented to the cultivation of innovative ability, the effect of personalized teaching is evaluated diversely through the use of a multivariate comprehensive evaluation that combines formative evaluation and summative evaluation. The formative evaluation of the process of cultivating innovation can be carried out through the way of portraying the portraits of individual students' and students' group's innovation ability again based on the evaluation index system of innovation ability, which can in turn guide and improve the problems in the teaching process.

(4) Teaching reflection

Teaching reflection is mainly based on the feedback results of multivariate teaching evaluation, which not only reflects the effect of cultivating innovative ability of individual students, but also reflects the teaching effect of the whole class, which is also an important guarantee to achieve the intended teaching goals, and also a need for teachers to formulate teaching strategies. Teaching reflection mainly focuses on the feedback results of student learning and the feedback results of teacher teaching. The feedback result of student learning refers to the difference between the current development status of students' innovative ability in various dimensions and the initial diagnosis of innovative ability in various dimensions in the preliminary stage of preparation for teaching. Based on this, the recommendation system will adjust the difficulty and direction of the educational resources recommended to the students; the students themselves can also clearly understand the current state of innovative ability and adjust their learning attitude and direction in a timely and targeted manner.

3. Quality assessment of talent cultivation model based on big data

3.1. BP Neural Network

The structure of the BP neural network [44] is shown in Fig. 4, which is capable of infinitely approximating the nonlinear function being solved.

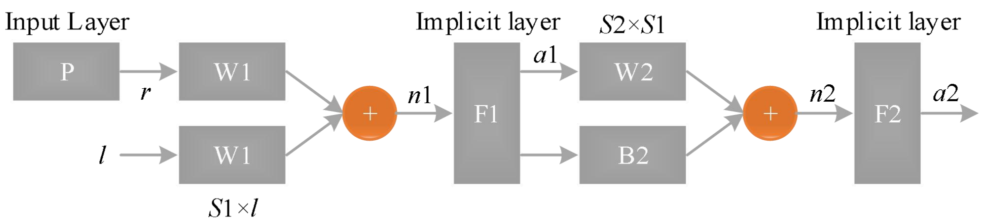


Figure 4. Basic structure of neural network.

Assuming that $x(t)$ and $y(t)$ are the input and output values of the samples, respectively, and that the function $f: R^m \to R^1$ is a mapping relation, where m denotes the input vector, the input values of the hidden layer of the BP neural network are able to be computed from equation (6).

$$S_j = \sum_{i=1}^{m} w_{ij} x_i - \theta_j \tag{6}$$

Here W_{ij} is the weights of the inputs and the implicit layer; θ_j is the threshold.

The Sigmoid function is chosen to construct the mapping function, which is defined as follows:

$$f(x)=\frac{1}{1+e^{-x}} \tag{7}$$

The output of the implicit layer is specified:

$$b_j=\frac{1}{1+\exp\left(\sum_{i=1}^{m}w_{ij}x_i-\theta_j\right)},j=1,2,\cdots,p \tag{8}$$

Similarly, the result of the output layer of the BP neural network can be calculated from equation (8).

$$\begin{cases} L=\sum\limits_{j=1}^{P}w_{jk}b_j-\partial_k \\ x_{i+1}=\dfrac{1}{1+\exp\left(\sum_{j=1}^{p}v_jb_j-\gamma\right)} \end{cases} \tag{9}$$

Where: v_j is the weights of the implicit and output layers; γ is the threshold.

3.2. Constructing a BP-based quality assessment model for talent cultivation

3.2.1. Establishment of a system of indicators for evaluating the quality of student training

Evaluation of student training quality is to analyze the factors affecting the quality of student training and the degree of influence through systematic collection of relevant information and adopting various qualitative and quantitative methods to arrive at the correct grade. Starting from the cultivation goal of vocational education, the index system of college student cultivation quality evaluation is divided into target level, first-level index, second-level index, and three levels as shown in Table 4. The target layer indicates the overall quality of the quality of training of college students; the first-level indicators are divided into knowledge, ability, and quality, and from the dynamic point of view, the formation of the quality of training is a gradual process, which can be divided into the process of mastering knowledge, the process of forming ability, and the process of developing quality. From a static point of view, there are three forms in a person at the same time: the form of knowledge, the form of ability, and the form of quality, and the three forms are combined in some form. The coordination of the three forms is constantly adjusted in the process of human development. The secondary indicators use measurable, accessible and comparable data indicators, which are the most basic elements in the indicator system, and can give direct or indirect measures to the quantity and intensity of the performance of the primary indicators, which are eventually transformed into quantitative values. The second-level indicator system can also continue to be refined, such as scientific and cultural knowledge, including basic knowledge of humanities and social sciences, basic knowledge of natural sciences, methodological knowledge, etc., and professional and technical knowledge can be divided into basic professional knowledge and professional knowledge. However, due to the limitations of statistical data, it is impossible to consider too many factors, of course, not the more the better, can only recognize the evaluation object from a limited point of view, but also taking into account the developmental nature of the students, the evaluation of the quality of student training is only limited to a certain stage, this paper is no longer divided into three-level indicators. Each evaluation index in the evaluation system of the quality of student training in colleges and universities is organically linked and works together to form an evaluation whole, which fully reflects the design idea of the index system.

Table 4. Quality evaluation index system of college students' talent training.

Target layer	Primary indicator	Secondary indicator
Quality of Education	Knowledge	Scientific and cultural knowledge
		Technical expertise
	Ability	professional ability
		Method capability
		Social skills
	Quality	Scientific and cultural literacy
		Professional competence

		Ideological and moral quality
		Physical and mental fitness

3.2.2. BP neural network talent cultivation quality evaluation model

According to the existing online education environment and the common ways of online learning, the evaluation model that can be used for the evaluation of students' learning behavior based on BP neural network is proposed.The steps for constructing the BP neural network talent cultivation quality evaluation model are generally as follows:

(1) Data collection. Data for training and evaluating the neural network need to be prepared first. These data can be user evaluations, scores of each evaluation and possibly other key factors, such as course name, course content, teacher competence and so on. At the same time, the initial sample data are collected from the student information system to determine the number of nodes in the input, implicit, and output layers as well as the expected value, taking into account the course learning of the students in our university.

(2) Data preprocessing. Preprocess the raw data, extract useful features, and standardize the data.

(3) Construct model. Construct the BP neural network model according to the evaluation index system.

(4) Model training. Train the constructed model to get the optimal parameters.

The parameter settings, the number of nodes in the input layer and the output are determined according to the specific problem, and there is no strict requirement for the number of hidden layers, which can be one or more layers. Kolmogorov's theory has proved that any given continuous function $\varphi : X - Y, X \in R^n, Y \in [0,1]^m, \varphi$ can be precisely realized by a three-layer neural network.

In addition to this, using too few neurons in the hidden layer may lead to an underfitting condition. Conversely, using too many neurons may result in overfitting, which in turn makes it difficult to achieve the desired results. Therefore, it is as important to pick the right number of neurons in the hidden layer. The number of neurons in the hidden layer L_h can be determined using the empirical formula (10), denoted as:

$$L_h = \frac{L_t}{\left(\left(L_i + L_0\right) * \beta\right)} \tag{10}$$

In expression (10), L_t is the sample data of the training set, L_i is the number of neurons in the input layer, L_0 is the number of neurons in the output layer, and β is an arbitrary variable, which is usually taken to be a constant in the interval [2,10]. The proposed activation function is a logistic function. The softmax function is then used to inverse normalize the output values of the system. The learning rate, training set and test set are determined by repeated experimental adjustments.

The BP algorithm is implemented and simulation experiments are carried out by applying SPSS software with the following steps:

Step 1: Normalization of input layer data using equation (11)

$$X = \frac{X - X_{\min}}{X_{\max} - X_{\min}} \tag{11}$$

Where X is the input data component, $X_{\min}$ and $X_{\max}$ denote the minimum and maximum values of this component in the training data set, respectively. After normalization of the input layer, the obtained values are in the interval $[0,1]$.

Step 2: Build the forward neural network and train it.

Step 3: Output the prediction results and verify the validity of the model, e.g., by cross-validation method.

Step 4: Apply the trained model to the evaluation of online teaching and get the online teaching quality evaluation results. Calculate the accuracy and reliability of the evaluation model and optimize the model.

3.3. Empirical analysis

Talent cultivation quality indicator data of three higher vocational colleges and universities from 2020 to 2024 are selected as the data for the empirical analysis of the assessment model to assess the quality of talent cultivation of three higher vocational colleges and universities during the period from 2020 to 2024. Using the BP assessment model that has been successfully tested and trained to input the talent cultivation indicator data of the higher vocational colleges and universities from 2020 to 2024 in Matlab, the simulation results of the talent cultivation quality indicator data of the higher vocational colleges and universities are realized by sim function to simulate the network simulation.

The fitting of the comparison between the predicted and actual values of the assessment of higher vocational colleges and universities is shown in Figure 5. It can be clearly seen that the trend of the predicted and real results of the assessment value of talent cultivation quality in higher vocational colleges and universities is very close to each other, with a consistent trend.

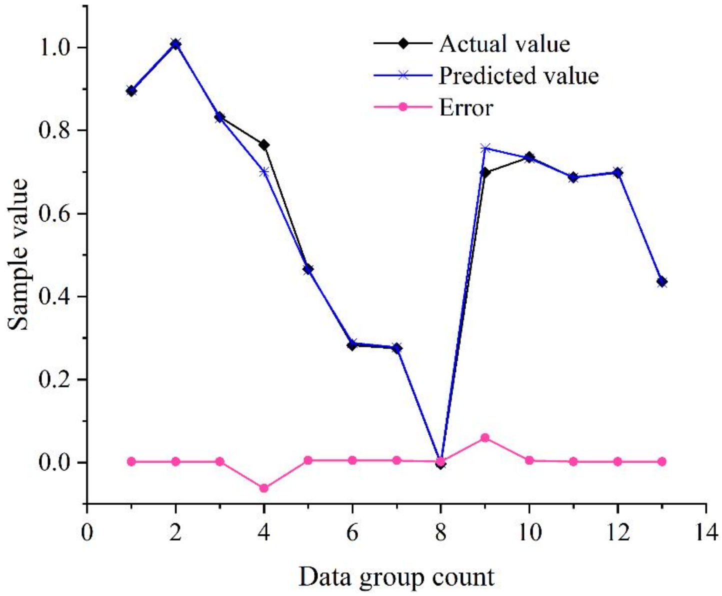


Figure 5. Diagram of talent training evaluation in higher vocational colleges.

The error change curve of the training process of the neural network model is shown in Fig. 6, and the model regression curve is shown in Fig. 7. It can be seen that only 3 epochs are used in the model running process to meet the desired requirements of this paper for the sum of error squares. In the model regression curve, the correlation coefficient is 0.99026, and the sample data are distributed on the diagonal, which indicates that the assessment model has a very high accuracy, and the results of the assessment grades of higher vocational colleges and universities from 2020 to 2024 are of great reference significance.

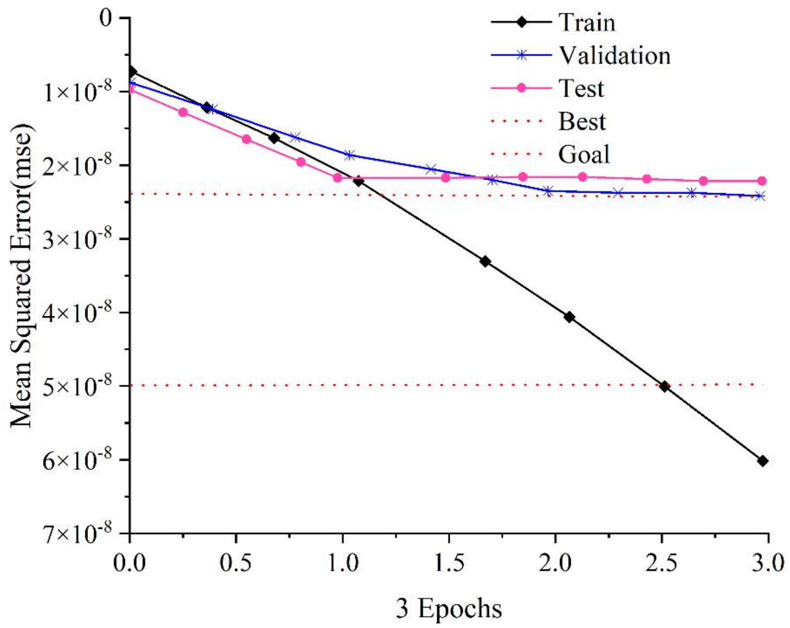
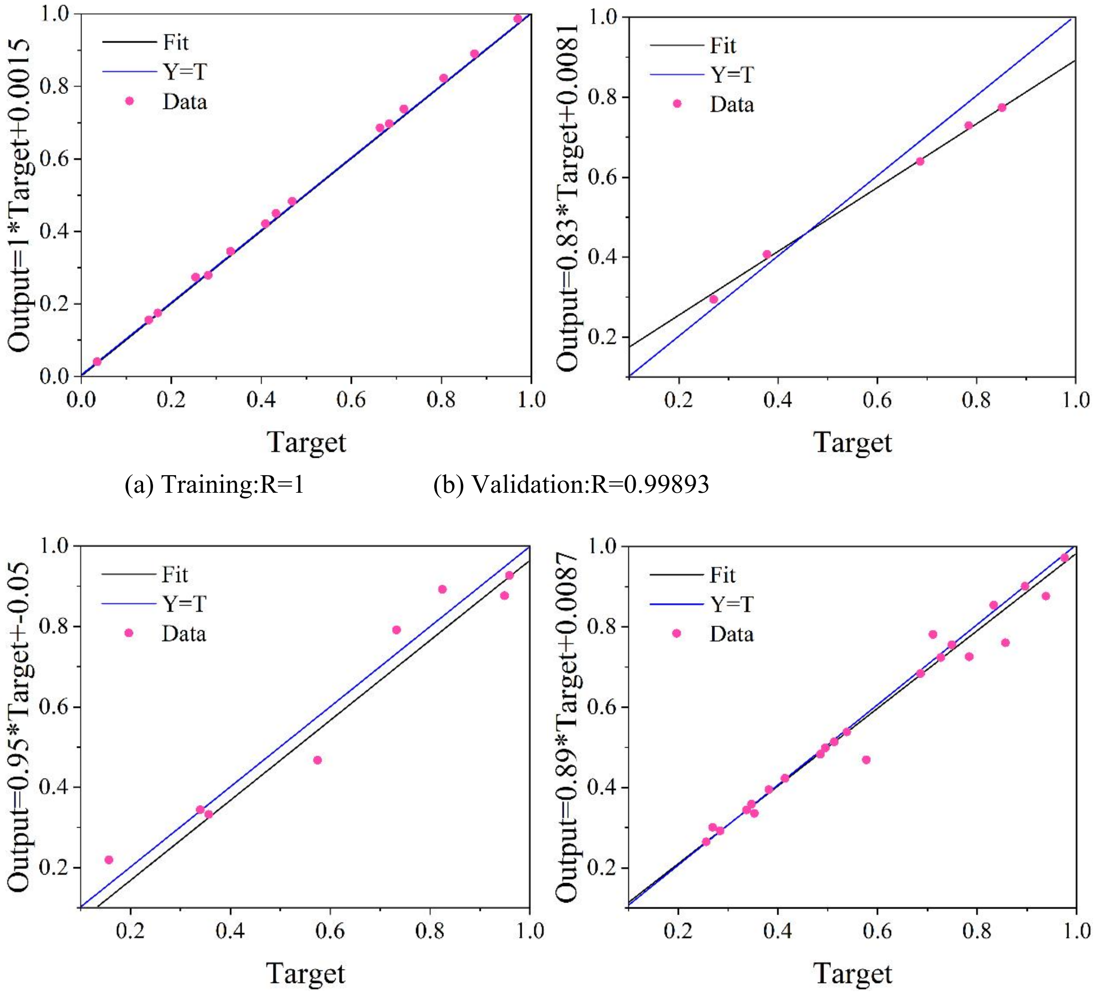


Figure 6. Error curve of talent training evaluation in higher vocational colleges.



(a) Training:R=1 (b) Validation:R=0.99893

(c) Teat:R=0.91153 (d) ALL:R=0.96374

Figure 7. Regression diagram of talent evaluation in colleges and universities.

The output results of the talent cultivation quality data of the three higher vocational colleges and universities from 2019 to 2024 take values in the range of [0,1], and the set of intervals is divided into {[1,0.8], (0.8,0.6], (0.6,0.4], (0.4,0.2], and (0.2,0]}, which correspond to the five grades of {excellent, good, good, qualified, and poor}, respectively. Therefore, after the empirical assessment of the BP talent training quality assessment model, the results of the assessment grades of the talent training quality of the three higher vocational colleges from 2020 to 2024 are obtained, as shown in Table 5.

As can be seen from the table, after the simulation of BP neural network model, the relative error between the predicted value and the real value of talent cultivation of the three higher vocational colleges and universities in the five years from 2020 to 2024 is within 7.7%, which indicates that the error is very small, and the model training accuracy is high, and the assessment results have a large degree of confidence. The results show that in the five-year period from 2020 to 2024, the quality of talent cultivation of H1 institutions has the best assessment grade, with three consecutive years of excellent assessment; the effect of the integration of industry and education of H2 institutions is the second best, with four consecutive years of better assessment grade and a stable upward trend; the talent cultivation of H3 institutions is the last, and there is one year of poor assessment grade. From this, it can be seen that the order of the assessment effect of industry-teaching integration from high to low is H1>H3>H1, and comprehensively, the average grade of the quality assessment of industry-teaching integration in higher vocational colleges and universities is good, and the mean value is 0.6100.

Table 5. Ranking table of talent training evaluation in higher vocational colleges.

College	H1 vocational college					H2 vocational college			H3 vocational college					Mean
Year	2024	2023	2022	2021	2020	2022	2021	2020	2024	2023	2022	2021	2020	
Predicted value	0.91	1.00	0.82	0.71	0.49	0.31	0.29	0.01	0.77	0.75	0.71	0.71	0.45	0.6100
Ground truth	0.91	1.00	0.83	0.78	0.49	0.31	0.29	0.01	0.72	0.75	0.71	0.72	0.45	0.6131
Absolute error	0.01	0.01	0.01	-0.07	0.01	0.01	0.01	0.01	0.01	0.01	0.01	-0.01	0.01	
Fractional error (%)	0.09	0.08	0.013	-8.59	0.25	0.19	0.76	0.01	0.01	0.01	0.01	-0.05	0.09	
Prediction Level	Outstanding	Outstanding	Outstanding	Preferably	Good	Qualified	Qualified	Range	Preferably	Preferably	Preferably	Preferably	Good	Good
Actual grade	Outstanding	Outstanding	Outstanding	Preferably	Good	Qualified	Qualified	Range	Preferably	Preferably	Preferably	Preferably	Good	Good

4. Conclusion

This paper proposes DEABC-FC clustering algorithm by constructing talent portraits and by introducing variation and crossover ideas in differential evolution, and the experimental results show that the DEABC-FC algorithm improves its solution efficiency and ability to jump out of the local optimum, which is conducive to the improvement of clustering effect. The DEABC-FC algorithm is used to perform clustering analysis on the student portraits, and design personalized teaching models for the three groups of students output from the student portraits, namely, self-disciplined schoolmaster type, night owl internet type and lack of planning type. Finally, based on the BP neural network, we constructed the index system for evaluating the quality of student talent cultivation and the talent quality evaluation model for empirical analysis, and the results show that the personalized teaching mode proposed in this paper is used for talent cultivation quality assessment effect from high to low in the order of H1>H3>H1, and comprehensively, the average grade of the quality assessment of the integration of industry and education in higher vocational colleges and universities is good, and the mean value is 0.6100.

References

1. Mascarenhas, C., Ferreira, J. J., & Marques, C. (2018). University–industry cooperation: A systematic literature review and research agenda. Science and Public Policy, 45(5), 708-718.
2. Figueiredo, N. L., & Ferreira, J. J. (2022). More than meets the partner: a systematic review and agenda for University–Industry cooperation. Management Review Quarterly, 72(1), 231-273.
3. Franco, M., & Haase, H. (2015). University–industry cooperation: Researchers' motivations and interaction channels. Journal of Engineering and technology Management, 36, 41-51.
4. Lopes, J., & Lussuamo, J. (2021). Barriers to university-industry cooperation in a developing region. Journal of the Knowledge Economy, 12(3), 1019-1035.

5. Yun, J., & Long, W. (2019, August). Thoughts on" Going Out" of School-enterprise Cooperation of Higher Education in Yunnan. In 5th International Conference on Arts, Design and Contemporary Education (ICADCE 2019) (pp. 1023-1027). Atlantis Press.
6. Zhang, L. (2016, April). The Collaborative Training Mechanism of Senior Talent Comprehensive Practice Teaching System of Administration of Innovation and Practice. In International Conference on Education, Management and Computing Technology (ICEMCT-16) (pp. 1430-1434). Atlantis Press.
7. Feng, F., Zhang, L., Du, Y., & Wang, W. (2015). Visualization and quantitative study in bibliographic databases: A case in the field of university–industry cooperation. Journal of Informetrics, 9(1), 118-134.
8. Kozlinska, I. (2012). Obstacles to university–industry cooperation in the domain of entrepreneurship. Journal of Business Management, 6.
9. Liu, X., & He, J. (2024). On personnel training of high-star hotel operations and management Major in secondary vocational schools by school-enterprise cooperation under the background of integration between industry and education. International journal of education and humanities, 12(3), 252-254.
10. Chen, H. (2022). Appraisal of Obvious Achievements and Hidden Values of School-Enterprise Cooperation Projects in vocational school. International Journal of New Developments in Education, 4(2), 22-27.
11. Li, Z. (2023). Research on school enterprise cooperation mode under the mechanism of "craftsman leading and collaborative education". In SHS Web of Conferences (Vol. 157, p. 02019). EDP Sciences.
12. Fei, C. H. E. N., Qi, X. I. E., & Huigang, X. U. (2016). Training Center of Application of Measurement and Control Technology Constructed by School-Enterprise Cooperation. Experiment Science and Technology, 14(2), 166-169.
13. Suh, Y., Woo, C., Koh, J., & Jeon, J. (2019). Analysing the satisfaction of university–industry cooperation efforts based on the Kano model: A Korean case. Technological Forecasting and Social Change, 148, 119740.
14. Yan, X., Zhang, W., Wang, H., Jia, L., Hou, Y., Xu, T., ... & Wang, Y. (2019, March). Construction of the Quality Assurance System for Collaborative Education of Schools and Enterprises Cooperation in Local Universities. In 2018 8th International Conference on Education and Management (ICEM 2018) (pp. 144-147). Atlantis Press.
15. Shao, J., Gao, C., & Zhou, L. (2023). Current Situation and Construction of Innovation Mechanism of Schoolenterprise Cooperation and Collaborative Education in Higher Vocational Education Based on Machine Learning. Adult and Higher Education, 5(11), 37-51.
16. Nam, G. M., Kim, D. G., & Choi, S. O. (2019). How resources of universities influence industry cooperation. Journal of Open Innovation: Technology, Market, and Complexity, 5(1), 9.
17. Awasthy, R., Flint, S., Sankarnarayana, R., & Jones, R. L. (2020). A framework to improve university–industry collaboration. Journal of Industry-University Collaboration, 2(1), 49-62.
18. Jiao, L., & Li, H. (2021). In the Era of Big Data, School Enterprise Cooperation, Production and Education Iintegration Build a Community of Common Destiny. In E3S Web of Conferences (Vol. 251, p. 01041). EDP Sciences.
19. Yu, B. (2023). Modern training model of apprenticeship based on multi-objective optimisation algorithm for sustainable development of school-enterprise cooperation. International Journal of Knowledge-Based Development, 13(2-4), 164-180.
20. Zhang, N. (2017, December). Analysis of the Cooperation Model and Problem of College and Enterprise. In 4th International Conference on Education, Management, Arts, Economics and Social Science (ICEMAESS 2017) (pp. 164-167). Atlantis Press.
21. Hui, X. (2025, May). Analysis and Research of the Personalized Practice Platform for School-Enterprise Integration Based on Big Data. In 2025 Global Conference in Emerging Technology (GINOTECH) (pp. 1-6). IEEE.
22. He, J., Lin, K. Y., & Dai, Y. (2022). A data-driven innovation model of big data digital learning and its empirical study. Information Dynamics and Applications, 1(1), 35-43.
23. Lin, H., Wan, S., Gan, W., Chen, J., & Chao, H. C. (2022, December). Metaverse in education: Vision, opportunities, and challenges. In 2022 IEEE International Conference on Big Data (Big Data) (pp. 2857-2866). IEEE.
24. Wang, Y., Wang, X., & Lin, Y. (2021, August). Innovation and practice of computer applied talents training mode under school enterprise cooperation education mode based on big data. In Journal of Physics: Conference Series (Vol. 1992, No. 3, p. 032100). IOP Publishing.
25. Qin, T. (2025). Innovation of interdisciplinary emergency language talent training model based on big data. Journal of Computational Methods in Sciences and Engineering, 14727978251385159.
26. HUANG, W., Qiu, N., Wang, F., Yan, C. Y., & Yi, W. T. (2019). Exploration and practice of school-enterprise collaborative education of applied talents. DEStech Transactions on Social Science, Education and Human Science,(icesd), 183-187.
27. Zhang, Y., & Yang, Y. (2021, May). Training Mode of School Enterprise Collaborative Innovation and Entrepreneurship Talents' Research. In 2021 2nd International Conference on Computers, Information Processing and Advanced Education (pp. 691-694).
28. Wan, J., & Hao, W. (2023). Talent Training Mode Under the Background of School Enterprise Cooperation and the Integration of Industry and Education. Journal of Theory and Practice of Humanities Science, 1(01), 31-36.
29. Xiao, P. P., & Zhang, Y. (2018, June). Research of how to cultivate the applied talents under school-enterprise collaboration. In 2018 2nd International Conference on Education, Economics and Management Research (ICEEMR 2018) (pp. 88-91). Atlantis Press.

30. Lihong, W., & Ping, L. (2020, March). Practical research on the talent training mode of school-enterprise cooperation based on the background of new engineering. In Proceedings of the 2020 International Conference on Education Development and Studies (pp. 22-26).
31. Peng, F., & Zhang, F. (2021). Research and Analysis on the Model of School-Enterprise Collaborative Education in Higher Vocational Colleges. Advances in Educational Technology and Psychology, 5(4), 66-72.
32. Zhou, C., You, L., Gao, Y., & Ning, F. (2019, October). Research and Practice on School-enterprise Collaborative Talent Cultivation Mode in Higher Vocational Colleges. In 2019 3rd International Conference on Economic Development and Education Management (ICEDEM 2019) (pp. 145-148). Atlantis Press.
33. Zhang, W., Wang, Z., & YifengCui. (2019, March). Realization of the enterprise value in university-enterprise cooperative talent-cultivating mode. In Proceedings of the 2019 7th International Conference on Information and Education Technology (pp. 200-203).
34. Huda, M., Anshari, M., Almunawar, M. N., Shahrill, M., Tan, A., Jaidin, J. H., & Masri, M. (2016). Innovative teaching in higher education: The big data approach. Tojet, 1210-1216.
35. Aithal, P. S., & Aithal, S. (2016). Ekalavya Model of Higher Education–An Innovation of IBM's Big Data University. International Journal of Current Research and Modern Education (IJCRME), 1(2), 190-205.
36. Chen, C., & Xu, W. (2020, April). Innovation and application of college students' education and management based on big data. In Proceedings of the 2020 3rd International Conference on Big Data and Education (pp. 5-9).
37. Xu, C., & Wang, S. (2022, July). Innovative Application of Big Data Technology in Network Teaching Model of University Courses. In International Conference on E-Learning, E-Education, and Online Training (pp. 578-588). Cham: Springer Nature Switzerland.
38. Huang, Y., & Jin, X. (2018). Innovative college English teaching modes based on big data. Kuram ve Uygulamada Egitim Bilimleri, 18(6), 3428-3434.
39. Wang, X. (2023). Research on Innovative Models of Accounting Professional Teaching in the Era of Big Data. Advances in Vocational and Technical Education, 2023, 5 (13).
40. Wang, C. (2022, January). Application of Big Data in the Innovation of Physical Education Teaching Mode. In Innovative Computing: Proceedings of the 4th International Conference on Innovative Computing (IC 2021) (pp. 639-646). Singapore: Springer Nature Singapore.
41. Yun, L. (2021, January). Innovation platform design of talent training mode of school and enterprise cooperation in higher vocational colleges under the background of big data. In 2021 13th International Conference on Measuring Technology and Mechatronics Automation (ICMTMA) (pp. 699-702). IEEE.
42. Xiaozhi Fan, Quanlai Li, Weipeng Zhang & Haonan Yin. (2025). Multi-Objective Process Parameter Optimization for Abrasive Air Jet Machining Using Artificial Bee Colony Algorithm. Machines,13(10),964-964.https://doi.org/10.3390/MACHINES13100964.
43. Anita Panwar & Satyasai Jagannath Nanda. (2024). Distributed Clustering in Wireless Sensor Network with Kernel Based Weighted Fuzzy C-Means Algorithm. SN Computer Science,5(8),1114-1114.https://doi.org/10.1007/S42979-024-03446-4.
44. Yanping Miao, Chiheng Cai, Xuhe Zheng, Changyue Liu, Jianxi Ren, Baojun Xu... & Jianqiang Yuan. (2025). Research on a coal-rock burst risk evaluation model based on particle swarm optimized BP neural network. Frontiers in Built Environment,11,1699662-1699662.https://doi.org/10.3389/FBUIL.2025.1699662.