Hindawi
Journal of Sensors
Volume 2023, Article ID 9837089, 1 page
https://doi.org/10.1155/2023/9837089

Retraction

Retracted: Analysis of Learning Ability of Ideological and Political Course Based on BP Neural Network and Improved *k*-Means Cluster Algorithm

Journal of Sensors

Received 8 August 2023; Accepted 8 August 2023; Published 9 August 2023

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1]. This investigation has uncovered evidence of one or more of the following indicators of systematic manipulation of the publication process:

(1) Discrepancies in scope
(2) Discrepancies in the description of the research reported
(3) Discrepancies between the availability of data and the research described
(4) Inappropriate citations
(5) Incoherent, meaningless and/or irrelevant content included in the article
(6) Peer-review manipulation

The presence of these indicators undermines our confidence in the integrity of the article's content and we cannot, therefore, vouch for its reliability. Please note that this notice is intended solely to alert readers that the content of this article is unreliable. We have not investigated whether authors were aware of or involved in the systematic manipulation of the publication process.

Wiley and Hindawi regrets that the usual quality checks did not identify these issues before publication and have since put additional measures in place to safeguard research integrity.

We wish to credit our own Research Integrity and Research Publishing teams and anonymous and named external researchers and research integrity experts for contributing to this investigation.

The corresponding author, as the representative of all authors, has been given the opportunity to register their agreement or disagreement to this retraction. We have kept a record of any response received.

References

[1] G. Zeng, "Analysis of Learning Ability of Ideological and Political Course Based on BP Neural Network and Improved *k*-Means Cluster Algorithm," *Journal of Sensors*, vol. 2022, Article ID 4397555, 11 pages, 2022.

Hindawi
Journal of Sensors
Volume 2022, Article ID 4397555, 11 pages
https://doi.org/10.1155/2022/4397555

Research Article

Analysis of Learning Ability of Ideological and Political Course Based on BP Neural Network and Improved k-Means Cluster Algorithm

Guidong Zeng

Xingjian College, Xijing University, Xi'an, Shaanxi 710123, China

Correspondence should be addressed to Guidong Zeng; zengguidong3@xijing.edu.cn

Received 4 January 2022; Revised 24 January 2022; Accepted 9 February 2022; Published 21 March 2022

Academic Editor: Yanqiong Li

With the rapid development of technologies such as big data analysis, machine learning, and cloud computing, artificial intelligence has made breakthrough progress in many fields. Artificial intelligence technology has also brought profound changes to higher education. Therefore, the ideological and political course in colleges and universities should integrate artificial intelligence technology into the teaching of ideological and political education and create an "intelligent ideological and political learning" to adapt to the goal of educational reform in the new era. This paper presents a research method of innovation ability of ideological and political course based on BP neural network and improved k-means clustering algorithm. Firstly, this method obtains the objective index that can comprehensively measure the learning ability through BP neural network and acquires the evaluation score of learning ability. Then, SPSS software is utilized to test the correlation between the influencing factors and the index, harvesting the factors that significantly affect graduate students' ideological and political learning ability. Finally, an improved k-means clustering algorithm is designed, which clusters the graduate students according to the different characteristics of the survey objects and gives targeted suggestions for each class of individuals to improve their ideological and political learning ability. The experimental results indicate that the proposed method is feasible and effective. The research method of ideological and political course ability proposed in this paper is of great significance to the promotion of ideological and political education in the era of big data.

1. Introduction

Artificial intelligence technology, as an important driving force of future educational reform, not only profoundly affects the traditional teaching mode of ideological and political education in colleges and universities but also poses a severe challenge to the orientation of the roles of teachers and students in the process of ideological and political education [1]. Based on the advantages of human-computer cooperation, cross-border integration, cocreation, and sharing brought by artificial intelligence technology, the personalized education model advocated by modern educational ideas has a practical basis.

Throughout the practice of educational informatization at home and abroad, the practical conditions for applying artificial intelligence technology to the field of education have matured, and related theoretical research and practical exploration are being carried out simultaneously [2]. Du et al. [3] proposed an English network teaching method based on artificial intelligence technology and WBIETS system, improving the deep learning network and taking it as the core algorithm of WBIETS system. Sun et al. [4] put forward the decision tree algorithm and the implementation model of English teaching evaluation based on neural network, which could help teachers improve their education level and students' English scores. The collaborative recommendation algorithm obtains a good accuracy in the course recommendation task according to the history of students' course selection records [5, 6]. In the big data environment, machine learning is used to predict learning results in online courses [7].

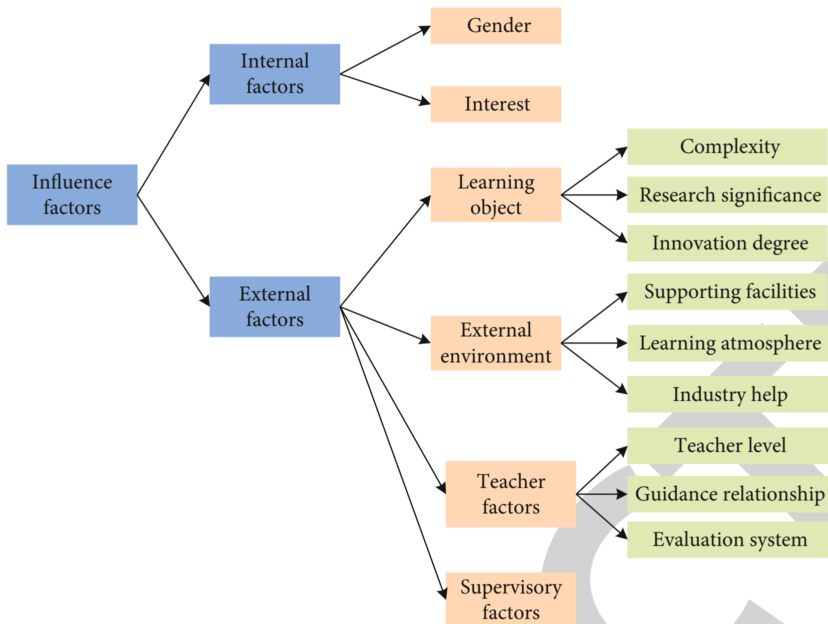

Figure 1: Related factors of graduate students' ideological and political learning ability.

Table 1: Statistical table of sample distribution ($n = 1200$).

Category	Project	Quantity	Percentage	Category	Project	Quantity	Percentage
Gender	Man	569	45.50%	Learning atmosphere	Good	527	42.40%
	Woman	681	54.50%		Common	723	57.80%
Interest	High	551	44.10%	Industry help	Enormous	310	24.80%
	Low	699	55.90%		More	321	25.70%
Degree of difficulty	Difficult	669	53.50%		Common	359	28.70%
	Easy	581	46.50%		Less	210	16.80%
Research significance	Great	434	34.70%		Minimum	50	4.00%
	Common	816	65.30%	Teacher level	Higher	916	73.30%
Degree of innovation	High	681	54.50%		Common	334	26.70%
	Low	569	45.50%	Learning relationship	Very harmonious	235	18.80%
Supporting facilities	Perfect	1166	85.30%		Fairly harmonious	644	51.50%
	Imperfect	184	14.70%		Common	297	23.80%
Supervisor	Timely and accurate	956	76.50%		Extremely discordant	74	5.90%
				Evaluation system	Perfect	1195	81.20%
	Need to be improved	294	23.50%		Imperfect	235	18.80%

The ideological and political theory course in colleges is the main channel to strengthen and improve the ideological and political education of students and postgraduates. In China, the demand and trend of the integration and innovation of artificial intelligence technology and ideological and political education are becoming more and more obvious, while there are still some bottlenecks in the integration process. Therefore, educational practitioners should reflect from the perspective of the gap between traditional teaching mode and modern technology, providing a three-dimensional thinking for the practice of "intelligent thinking and politics" in colleges and universities in China.

Educational data mining and learning analysis are new research fields. It is worth studying if various statistical and machine learning methods are applied to enhance ideological and political classroom education. Neural network and clustering algorithm are commonly used data mining methods. Neural network is widely used in pattern recognition, analysis, control, and prediction. Literature [8] proposes research on the optimization of scientific research performance management based on BP (back propagation) neural network. This algorithm uses neural network to construct the performance evaluation model of social science research in colleges and universities. The experimental results demonstrate that the model is an efficient evaluation method. BP neural network extracts six paper features, two journal features, nine author features, eight reference

Table 2: Evaluation score of graduate students' ideological and political learning ability.

Evaluation score of ideological and political learning ability			Theoretical achievement		
			Be poor	Common	Good
Participation in ideological and political ability competition	Participated in	Win a prize	3	6	9
		Unawarded	2	5	8
	Did not attend		1	4	7

Table 3: Statistical table of correlation and difference between postgraduate's learning ability and its internal factors.

Project	Relevance P	Difference S_{ig}
Gender	0.283	0.161
Interest	0.007	0.003

features, and five early citation features to predict the citation times of a single paper [9].

In the era of big data, clustering analysis of massive data is an important research direction. Clustering algorithm has been widely used in education, e-commerce, transportation, and other fields [10]. k-means is widely used because of its high efficiency and easy understanding. However, the initial clustering center of traditional k-means algorithm is randomly selected, which easily leads to the clustering result falling into local optimum [11]. Meanwhile, the random method will also lead to the instability of the initial clustering center selection, which makes the clustering result unstable. Many scholars have done a great number of researches on the initial clustering center selection of k-means algorithm. For example, adaptive cuckoo and gravity search algorithm are used to optimize the initial cluster center by introducing swarm intelligence algorithm. However, the algorithm has not been applied because of its complexity [12]. Some scholars optimize the initial clustering center from the perspective of sample data density and distance. For example, Kalevala et al. [13] considered local distance to optimize the algorithm. Tang et al. [14] consider density and distance step by step, but calculating the data weight increases the time consumption. Yu et al. [15] proposed LOF algorithm to build a potential background dictionary from the perspective of local density and effectively excluded abnormal objects by calculating local density and abnormal values, yet there was a problem of inaccurate selection of cluster centers.

Domestic colleges and universities are also paying more and more attention to the cultivation of ideological and political classroom thinking ability and innovation ability of college students. Nevertheless, the effect is not obvious. This paper studies how to improve the ideological and political learning ability of postgraduates. Firstly, the influencing factors of learning ability are preliminarily determined through data collection and screening of effective information. Then, a specific standard to measure the abstract term learning ability is formulated according to BP neural network. Thus, the correlation between learning ability and influencing factors is analyzed to screen out the influencing factors with significant correlation. Finally, these factors will be used as variables to improve k-means clustering, and specific suggestions will be put forward for each type of individuals.

2. Algorithm Model in This Paper

2.1. Collect Data to Preliminarily Determine the Influencing Factors. This paper investigates the factors that affect students' learning ability in ideological and political class. Data were collected by literature survey, questionnaire survey, and focus interview. The interview outline is listed on the basis of literature survey. Graduate students and graduate tutors from different majors were invited in the form of interview groups to deeply explore the internal and external factors related to learning ability. The internal factors affecting learning ability include gender and interest, as shown in Figure 1.

External factors include learning objects, external environment, teacher factors, and supervision factors. The learning object is subdivided into the difficulty, innovation, and research significance of the subject. The external learning environment is subdivided into teaching facilities, learning atmosphere, and ideological and political help to the industry. Teachers' factors are subdivided into teachers' level, guiding-learning relationship, and evaluation system. Based on this, a questionnaire was compiled. In this study, 1230 formal questionnaires were distributed, and 1200 were effectively recovered. The effective questionnaire recovery rate was 97.56%, which was statistically significant. Sample distribution is shown in Table 1.

2.2. Evaluation Model of Students' Learning Ability Based on BP Neural Network

2.2.1. BP Network Design. The design of BP network includes the input layer, the output layer, the number of nodes in the hidden layer, and the transfer function between layers.

(1) Enter the Number of Layer Nodes. The number of input layer nodes corresponds to the number of evaluation indexes. Based on many research findings, the evaluation indexes are test scores, creative ability, scientific research ability, paper writing ability, and competition level. According to the analysis, the evaluation indexes of graduate students' ideological and political learning ability are as

Table 4: Statistical table of correlation and difference between postgraduates' learning ability and external factors.

Project		Relevance P	Difference S_{ig}
	Difficulty degree of subject	0.01	0.015
Learning object	Research significance	0.131	0.253
	Degree of innovation	0.009	0.02
	Learning atmosphere	0.003	0.084
External environment	Teaching facilities	0.878	0.68
	Industry help	0	0.005
	Teacher level	0.01	0.014
Teacher factor	Learning relationship	0.009	0.012
	Evaluation system	0.007	0.013
Supervision factor	Supervising work	0.005	0.006

follows: $5\,(x1-x5)$, so these five evaluation indexes are taken as input nodes $n=5$.

(2) Number of Output Layer Nodes. This paper takes the final evaluation result as the output of the network. Number of output nodes $m=1$.

(3) Number of Hidden Layer Nodes. Based on the Kolmogorov theorem proved by Hecht-Nielsen, three-layer BP neural network can approximate any continuous function under reasonable structure and proper weight conditions. Therefore, the three-layer BP network is selected in order to simplify the calculation.

There is no optimal theoretical method to determine the number of hidden layer nodes, which is a more complicated problem. Too few nodes will lead to poor fault tolerance. Too much network training time will be increased and generalization ability will be reduced. Therefore, the designer's experience and many experiments are usually used to determine the optimal number of hidden nodes. This paper chooses the number of implicit nodes $s=3$ after empirical analysis.

2.2.2. BP Network Learning Algorithm

(1) Input: dataset D. Learning rate α: $\alpha \in [0,1]$. Stop condition: the error rate specifies the threshold θ. The maximum number of iterations is $T*$

(2) Initial link weight: $T=0$, $g_{pq}^{(T)}$, $s_q^{(T)}$, $\omega_{qk}^{(T)}$, $b_k^{(T)}$

(3) Input samples (u_h, x_h) in turn, and calculate the expected predicted value x_k

(4) Update the link weight:

$$\begin{aligned} g_{pq}^{(T+1)} &= g_{pq}^{(T)} + \Delta g_{pq}^{(T)}, \\ s_q^{(T+1)} &= s_q^{(T)} + \Delta s_q^{(T)}, \\ \omega_{qk}^{(T+1)} &= \omega_{qk}^{(T)} + \Delta \omega_{qk}^{(T)}, \\ b_k^{(T+1)} &= b_k^{(T)} + \Delta b_k^{(T)}, \end{aligned} \tag{1}$$

where $g_{pq}^{(T)}, \omega_{qk}^{(T)}$ represents the forward propagation connection weight. $s_q^{(T)}$, $b_k^{(T)}$ represents the back propagation connection weight

(5) $T=T+1$, judge whether the stopping condition is met. If the model error is less than the specified threshold or the maximum iteration times are greater than the threshold, stop the iteration. Otherwise, return to Step 3

(6) Output: $g_{pq}^{(T)}$, $s_q^{(T)}$, $\omega_{qk}^{(T)}$, $b_k^{(T)}$. The output value of the kth neuron in the output layer is

$$\widehat{x}_k = f\left(\sum_{q=1}^{d} \omega_{qk} z_q + b_k\right), \tag{2}$$

where z_q indicates hidden input

2.2.3. Application of BP Network Model. Based on the above analysis, this paper uses Python to build a three-layer BP neural network with 5 input neurons, 3 hidden layer neurons, and 1 output neuron. When the five evaluation indexes: examination score, creation ability, scientific research ability, thesis writing ability, and competition level, are used as inputs, the input data (training samples) need to be normalized. The linear function is used as the transfer function because the input layer only transmits data. The neurons in the hidden layer adopt S-type function (Sigmoid function). At the same time, the learning rate is 0.6, and the convergence error threshold is 0.001. Finally, the evaluation results were reverse-normalized and the evaluation scores were obtained as shown in Table 2.

The number of training samples selected is 200, which can satisfy the fitting. 100 people are employed to predict the test model. The prediction results are compared with the expert evaluation results, and the prediction accuracy is observed. Finally, it demonstrates that the evaluation model of graduate students' ideological and political learning ability based on BP neural network is reasonable and effective.

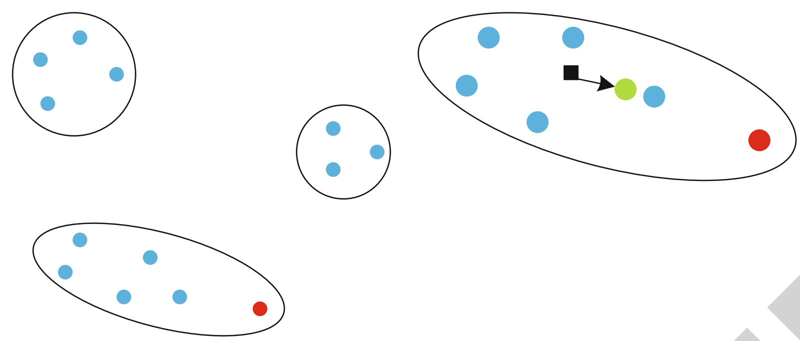

FIGURE 2: Replacement of outliers and pseudocenters.

2.3. Screening the Significant Influencing Factors of Ideological and Political Learning Ability. The influencing factors of graduate students' ideological and political learning ability collected by qualitative investigation methods such as focus interview and expert discussion are subjective and can only be defined as a preliminary scope. This paper studies the correlation degree between these factors and ideological and political learning ability in quantitative form after determining the measurement standard according to BP neural network. In this way, the factors with significant correlation can be screened out, and the weak correlation factors can be filtered out, ensuring the accuracy of the research. Based on the evaluation score of ideological and political learning ability obtained by the BP neural network mentioned above, the correlation degree between the internal and external influencing factors such as ideological and political learning ability and interest, learning atmosphere, evaluation system, teacher's teaching, and ideological and political learning ability is studied. In this study, SPSS 22.0 was used for Spearman correlation analysis.

The Spearman correlation coefficient between two random variables u and x is recorded as r [16], and its formula is

$$r = 1 - \frac{\left(6\sum_{p=1}^{n} D_p^2\right)}{\left[n\left(n^2-1\right)\right]}, \quad (3)$$

where

$$\sum_{i=1}^{n} D_p^2 = \sum_{p=1}^{n} \left(U_p - V_p\right)^2, \quad (4)$$

where u represents the evaluation score of learning ability. x represents 9 significant influencing factors. U_p and V_p represent the rank after sorting variables u and x, respectively. n represents the sample size.

2.3.1. Research on Internal Influencing Factors. Table 3 is a statistical table of the correlation and difference between postgraduate's learning ability and internal influencing factors. According to the correlation test, gender and interest are significantly related to learning ability ($P < 0.01$). According to the difference test, there are significant differences between the two variables and learning ability ($P < 0.05$).

There is no significant correlation between gender and learning ability ($P > 0.01$), and there is no significant difference between gender and learning ability ($P > 0.05$).

Interest and ideological and political learning ability have a very significant correlation ($P < 0.01$), and the difference between interest and learning ability is extremely significant ($P < 0.05$). Among them, the more interest in learning, the stronger the learning ability, indicating a positive correlation trend. This phenomenon also accords with people's consistent thinking. The more interested you are, the more time and energy you put into it, and the stronger your ideological and political thinking ability.

2.3.2. Research on External Influencing Factors. Table 4 is a statistical table of the correlation and difference between graduate students' ideological and political learning ability and external influencing factors. The correlation test indicates that the difficulty and innovation degree of the subject in the learning object, the learning atmosphere and industry help in the external environment, all variables in the teacher factors and supervision work have significant correlation with the ideological and political learning ability ($P < 0.01$) and significant difference ($P < 0.05$). The correlation and difference of other variables are not significant.

Learning object: the difficulty and innovation degree of the subject in the learning object have a very significant correlation with the learning ability ($P < 0.01$), and the difference in the influence on the ideological and political learning ability is extremely significant ($P < 0.05$). However, the relevance and difference of research significance are not notable. Among them, the higher the learning ability, the higher the difficulty and innovation degree of the subject. This just indicates that people with strong ideological and political learning ability have stronger ability to solve difficult problems, while the significance of research is the selectivity of topics, which has little to do with learning ability.

The external environment: the learning atmosphere and industry help in the external environment have a very significant correlation with the ideological and political learning ability ($P < 0.01$), and the difference in their influence is extremely significant ($P < 0.05$). However, the relevance and difference of teaching facilities are not significant. Among them, the study of ideological and political courses has a great relationship with the learning atmosphere of the school. The stronger the learning atmosphere, the higher the enthusiasm for learning ideological and political affairs. If the future industry needs more ideological and political education, the demanders will study more enthusiastically.

All the variables in teachers' factors have a very significant correlation with the ability of ideological and political learning ($P < 0.01$), and they have a significant influence on the ability of ideological and political learning.

The difference of influence is extremely enormous ($P < 0.05$). Among them, the higher the level of teachers and the more concerned about students, the higher the students' learning enthusiasm. The stricter the evaluation system is, the harder students will study in order to pass the exam, which is positively related to their learning ability.

Factors of supervision: there is a very significant correlation between supervision and ideological and political learning ability ($P < 0.01$), and the difference of influence on ideological and political learning ability is extremely remarkable ($P < 0.05$). Among them, the timelier the teacher's supervision is done, the students will naturally keep up with the learning progress in time.

In this study, questionnaire survey and group discussion were used to explore the ideological and political learning ability and influencing factors of postgraduates, and the following conclusions were drawn. (1) Both internal and external factors of ideological and political learning ability had significant influence on it. Among them, the internal factors include interest, and the external factors include the difficulty and innovation of the subject in the learning object, the learning atmosphere, and industry help in the external environment, all variables in the teacher factors and supervision. (2) Interest and industry help have the greatest correlation with ideological and political study.

2.4. Improve the Clustering of Significant Influencing Factors of k-Means Algorithm. In order to improve the clustering accuracy of significant influencing factors of ideological and political ability, this paper proposes an improved k-means algorithm DC k-means (density parameter and center replacement k-means) based on density parameters and center replacement. The algorithm uses the density parameters of data objects to gradually determine the initial cluster center and uses the center replacement method to update the initial center that deviates from the actual position. Therefore, DC k-means is more accurate than the traditional clustering algorithm.

The DC k-means proposed in this paper firstly determines the initial cluster center by calculating the density parameters of each data object in the dataset, avoiding the unstable clustering result caused by randomly selecting the initial cluster center. Secondly, the biased cluster centers generated by the traditional k-means algorithm are replaced to avoid the influence of outliers on the clustering results.

2.4.1. Selection of Initial Cluster Centers Based on Density Parameters. DC k-means adopts the strategy of selecting

Table 5: Software and hardware configuration environment of the experiment.

CPU	Inter(R) Core (TM) i7-8565U CPU @ 1.80 GHz
RAM	LPDDR3 2133 MHz (8 GB)
Hard disk	NVMe PCIe high-speed solid-state drive
OS	Microsoft Windows 10 Enterprise (64 bit)

cluster center based on density parameter increment. This section and subsequent discussions assume that in Euclidean space R^m, dataset $D = \{y_1, y_2, \cdots, y_n\}$ contains n data objects. Every object $y_i = \{y_{p1}, y_{p2}, \cdots, y_{pm}\}$ has m attributes. Dataset D is divided into k clusters by a clustering algorithm $C = \{C_1, C_2, \cdots, C_K\}$, where $|C_K|$ is the number of data objects contained in the class cluster C_K. The corresponding center point of each cluster in cluster set C is $V = \{V_1, V_2, \cdots, V_K\}$. Euclidean distance $d(y_p, y_q)$ between any two data objects y_p and y_q in dataset D is defined as

$$d\left(y_p, y_q\right) = \sqrt{\left(y_{p1} - y_{q1}\right)^2 + \left(y_{p2} - y_{q2}\right)^2 + \cdots + \left(y_{pm} - y_{qm}\right)^2}. \quad (5)$$

Based on Euclidean distance, the maximum distance (LaDist) and minimum distance (SmDist) between all data objects are defined as follows:

$$\begin{aligned} \text{Laipst} &= \sum_{p=1}^{n-1} \max_{1 \le p < q \le n} d\left(y_p, y_q\right)^2, \\ \text{Smipst} &= \sum_{p=1}^{n-1} \min_{1 \le p < q \le n} d\left(y_p, y_q\right)^2. \end{aligned} \quad (6)$$

Although it is assumed above that the dataset D is divided into k clusters, the number of data objects in each cluster generated by different clustering algorithms may be different. As the number of data objects changes, the distance between each data object pair will also alter. Define the dynamic average distance (Divests) based on the maximum distance and minimum distance between all data objects:

$$\frac{\text{Divests} = (\text{LaDist} + \text{SmDist})}{(2 * K)}, \quad (7)$$

where k is the number of clusters into which dataset D is divided. According to the dynamic average distance, the density parameter can be defined as follows.

In dataset D, the number of data objects in the circular area with y_p as the center and Divests as the radius is called

Table 6: Description of 8 simulated datasets and 8 UCI real machine datasets.

Datasets	Points number	Cluster number	Dimension
Normal	200	5	2
D900	900	9	2
R15	600	15	2
N7	28000	7	2
K3	102000	3	2
Curve	180	3	2
Pathbased	300	3	2
Semicircle	300	3	2
Iris	150	3	4
Seeds	210	3	7
Haberman	306	2	3
Column	310	3	6
Hayes-Roth	132	3	5
Ionosphere	351	2	34
PageB locks	5473	5	10
Magic	19020	2	10

Table 7: Comparison of precision purity of different algorithms.

Datasets	k-medoids	k-means++	DC k-means
Normal	68.52	82.57	99.5
D900	75.29	81.13	99.78
R15	73.22	92.24	99.67
N7	82.3	88.16	100
K3	86.9	86.9	100
Curve	67.5	89.91	99.22
Pathbased	61.42	75.27	75.25
Semicircle	88.05	100	100
Iris	77.09	83.22	92.67
Seeds	69.64	76.55	89.12
Haberman	72.92	73.31	86.97
Column	70.41	72.25	72.31
Hayes-Roth	43.37	47.99	47.99
Ionosphere	67.12	71.33	72.53
PagcBIocks	91.03	90.99	92.8
Magic	65.89	66.09	66.48

the density parameter of data object y_p, that is,

$$\rho\left(y_p, \mathrm{DyAveDpst}\right) = \sum_{p=1,q\neq p}^{n} j\left(\mathrm{DyAveDpst} - d\left(y_p, y_q\right)\right), \tag{8}$$

where $j()$ is a jump function. $j(y) = 1$ when $y \geq = 0$; otherwise, $j(y) = 0$.

In the process of finding the center of initial cluster, most density-based clustering algorithms are more dependent on external parameters in the choice of neighborhood radius. Improper parameter selection will greatly affect the performance of the algorithm. To solve this problem, this paper first defines the dynamic average distance (Divests) based on the maximum distance and minimum distance between all data objects. The distance dynamically changes with each iteration, which can obtain the neighborhood radius of different division stages in time, to determine the density parameters more efficiently and stably and reduce the influence of external parameters on clustering results.

2.4.2. Replacement of Cluster Center. Another defect of traditional k-means algorithm is that it is very sensitive to outliers of dataset. In fact, some initial cluster centers generated by traditional k-means algorithm are not real cluster centers in the target dataset (this paper calls these points "pseudocenters"). In addition, the position of the generated cluster center will deviate from that of the actual cluster center due to the influence of outliers. This problem will seriously reduce the accuracy of traditional k-means algorithm.

The center generated by k-medoids clustering algorithm is always the real data point of the target dataset. Inspired by k-medoids algorithm, this paper proposes to update the pseudocenters generated by traditional k-means algorithm by using center replacement method. Once the k-means algorithm creates a pseudocenter for a class cluster, it will be replaced by the nearest point in the class cluster. At the same time, the neighboring point should be as far away from the outliers of this cluster as possible. In the process of clustering, the pseudocluster centers are updated in turn until all the real cluster centers are specified.

Figure 2 shows the replacement process of clusters containing outliers and their corresponding pseudocenters. In Figure 2, the blue dots represent normal data objects and the red dots represent outliers. Figure 2(a) shows a dataset composed of three clusters randomly generated by Python software. The cluster at the bottom left of Figure 2(a) contains an outlier represented by a red dot. Figure 2(b) shows the replacement process of the center point of this kind of cluster. Without outlier interference, the traditional k -means algorithm takes the object represented by the black rectangle as the cluster center. However, if the outliers in the cluster are considered, the obtained cluster center will deviate from the "actual" cluster center. As shown in Figure 2(b), along the arrow direction, the center of the cluster moves from the position of the black rectangle to the position of the green rectangle. This deviation will lead to the performance degradation of clustering algorithm. In fact, with the deviation of cluster center, many data objects that do not belong to this cluster will be included in the next iteration of clustering algorithm. In Figure 2(b), this paper uses an improved method to take the blue dot in the red rectangular box as the final cluster center. The center is the actual data object, which is closest to the black rectangle and as far away from the red outlier as possible.

2.4.3. DC k-Means Process and Time Analysis. The flow of DC k-means algorithm is shown in Algorithm 1. DC k -means algorithm can not only find the center of initial cluster stably but also can deal with outliers. In Algorithm 1: (1)

Table 8: Summary of correlation of influencing factors.

Project	Interest	Teacher level	Degree of innovation
P value	0.007	0.01	0.009
Project	Gender	Learning relationship	Evaluation system
P value	0.283	0.009	0.007
Project	Research significance	Teaching facilities	Learning atmosphere
P value	0.131	0.878	0.003
Project	Supervising work	Difficulty degree of subject	Industry help
P value	0.005	0.01	0

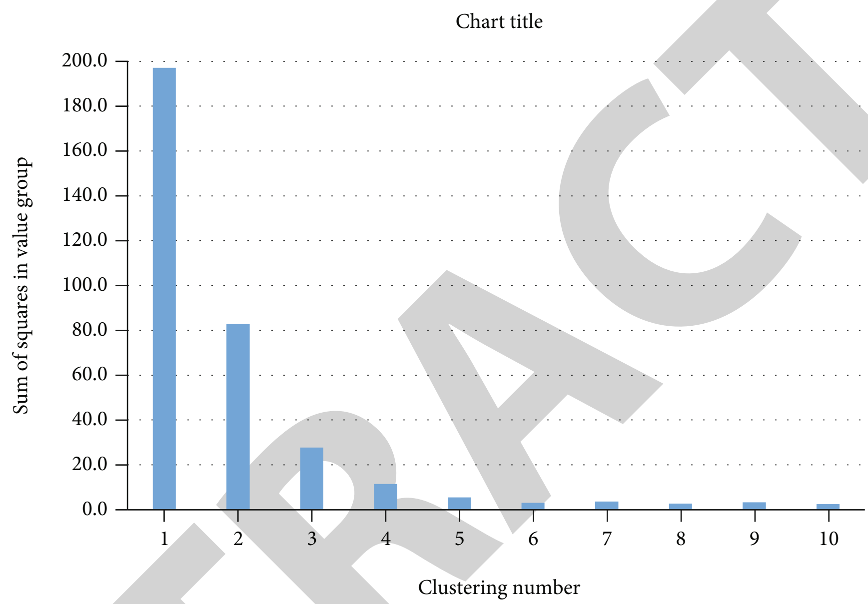

Figure 3: Sum of squares in value group under different clustering numbers.

Table 9: Final cluster center.

	Cluster 1	2	3	4	5
Learning ability evaluation score	7	8	1	4	1
Interest	High	High	Low	Low	Low
Learning relationship	Very harmonious	General harmony	Common	Common	General harmony
Industry help	Common	Enormous	Smaller	Larger	Enormous
Difficulty degree of subject	Easy	Difficult	Easy	Difficult	Difficult
Degree of innovation	Low	High	Low	Low	High
Supervision factor	Accurate and timely	Accurate and timely	Accurate and timely	Need to be improved	Need to be improved
Learning atmosphere	Good	Good	Common	Good	Common
Teacher level	Higher	Common	Higher	Higher	Higher
Evaluation system	Imperfection	Improve	Improve	Improve	Improve

calculate the dynamic average distance between all data objects in dataset D. (2) Calculate the density parameters of all data objects. (3) Find k initial cluster centers of dataset D and put them into set V according to the density parameter. Steps (4)–(8) realize the final division of dataset D. Specifically, step (5) initializes each cluster. (6) Put the data object into the corresponding class cluster. (7) Update the cluster center by using the center replacement method.

Input: dataset $D = \{y_1, y_2, \cdots, y_n\}$, number of clusters K.
Output: $C = \{C_1, C_2, \cdots, C_K\}$ of dataset D.
Algorithm:

(1) Calculate the dynamic average distance (Divests) between any pair of data objects (y_p, y_q) in dataset D

(2) For $p = 1, 2, \cdots, n$, do calculate the density parameter $\rho(y_p, \text{DyAveDpst})$ of the data object y_p

(3) For $k = 1, 2, \cdots, K$, do//find k initial cluster centers and put them into the initial cluster center set V. Select the data object y with the highest density parameter from dataset D, and delete all data objects about y from dataset D. Set y as the center of the kth initial cluster and v_k. $V \longleftarrow v_k$;//put v_k into set V in the center of the initial cluster

(4) Repeat

(5) Let $C_k = \varnothing (1 \leq k \leq K)$//initialize each cluster C_k

(6) For $p = 1, 2 \cdots, n$ do//put each data object in D into the corresponding cluster. Calculate the distance between the centers of each cluster in data objects y_p and V. Put y_p into the corresponding class cluster according to the nearest principle

(7) For $k = 1, 2, \cdots, K$, do//update the center of each cluster. Calculate the distance between the center v_k of the cluster C_k and other data objects in the cluster; find the nearest data object $(v_k{}')$ to v_k. At the same time, $(v_k{}')$ should be as far away from the outliers in C_k as possible. If $v_k \neq v_k{}'$ $v_k \longleftarrow v_k{}'//v_k{}'$ is updated as the new center of class C_k

(8) Until $\sum_{i=1}^{K} \sum_{X \in C_i} d(v_i, x)^2$ converted//until the standard function $\sum_{i=1}^{K} \sum_{X \in C_i} d(v_i, x)^2$ converges to a constant, at which point v_k is the new center of class C_k

Assume that there are n data objects in dataset $D = \{y_1, y_2, \cdots, y_n\}$. Every data object $y_i = \{y_{i1}, y_{i2}, \cdots, y_{im}\}$ is a dimension vector of m. The DC k-means algorithm will divide dataset D into k clusters $C = \{C_1, C_2, \cdots, C_K\}$ through h iterations.

According to the individual's theoretical and practical ability, BP neural network divides the learning ability of ideological and political education into nine levels. Subjective and abstract indicators are standardized, making individuals more different. It is found that the factors significantly related to the evaluation scores of ideological and political learning abilities are interest, guidance and learning relationship, industry help, project difficulty, innovation degree, teachers' level, supervision factors, evaluation system, and learning atmosphere. Those factors that have significant correlation and difference with the ideological and political learning ability and the evaluation score of the ideological and political learning ability are selected as cluster variables so that there are significant differences among all kinds, and the significance of cluster analysis is clarified. In the experimental part, DC k-means cluster analysis will be used to classify different research individuals, and according to the characteristics of each category, targeted suggestions will be put forward for their graduate students to improve their ideological and political learning ability.

3. Experiment and Analysis

3.1. Experimental Configuration and Indicators. The experimental code is written by Python language, aiming to verify the actual effect of this algorithm in the analysis of learning ability of ideological and political courses in colleges and universities. The specific software and hardware configuration environment of the experiment is listed in Table 5.

The experimental dataset for verifying DC k-means clustering algorithm consists of 8 simulation datasets and 8 UCI real machine learning datasets, as shown in Table 6. The data used to verify the ideological and political ability analysis algorithm in this paper consists of real student data collected in Section 2.1.

3.2. Performance Indicators. The accuracy of clustering results can usually be measured by external evaluation indexes, such as F-measure, entropy, and purity. This paper uses purity index to evaluate the accuracy of clustering results, which is defined as

$$\left(\text{purity} = \sum_{p=1}^{K} \frac{|C_p|}{n} \text{may} \left(\frac{m_{pq}}{|C_p|} \right) \right), \tag{9}$$

where $|C_p|$ is the number of all data objects in the cluster C_p. m_{pq} is the number of members of class cluster C_p belonging to class cluster C_q. K is the number of clusters in the target dataset. n is the number of data objects contained in the target dataset. In this paper, the value of purity index is converted into percentage for comparison.

3.3. Comparative Experiment. In the experiment, the clustering results generated by different datasets are evaluated.

Table 7 lists the processing accuracy of k-medoids, k-means++, and DC k-means for the 16 datasets listed in Table 6. The DC k-means algorithm can keep the same result every time, so it only needs to be run once. On the contrary, the accuracy of the other two algorithms is the average of 10 repeated experiments. As can be seen from Table 6, since the center of the initial cluster is randomly selected, the accuracy of k-medoids algorithm is the worst among the three algorithms. In k-means++ algorithm, except for the first initial cluster center, other cluster centers are no longer randomly selected. Therefore, the accuracy of k-means++ algorithm is better than that of k-medoids algorithm. DC k-means introduces the density parameter to select the initial cluster center and adopts the center replacement strategy in the update stage. Therefore, the clustering accuracy of DC k-means is the best among the three clustering algorithms, indicating the effectiveness of the improved k -means clustering algorithm designed in this paper.

3.4. Feasibility Analysis. According to the individual's theoretical and practical ability, the learning ability is divided into 9 levels by BP neural network, which standardizes subjective and abstract indicators, thus making individuals more different. Through correlation analysis, it is found that the factors significantly related to the evaluation score of ideological and political learning ability are interest, guidance and learning relationship, industry help, project difficulty, innovation degree, teachers' level, supervision factors, evaluation system, and learning atmosphere. These factors with significant correlation and difference with ideological and political learning ability and the evaluation score of ideological and political learning ability are selected as cluster variables to make significant differences between various types, so as to clarify the significance of cluster analysis. The following uses the improved cluster analysis method to classify different research individuals and puts forward targeted suggestions for their graduate students to improve their ideological and political learning ability according to the characteristics of each category.

According to Table 8, among the preliminarily determined influencing factors, only gender, research significance, and teaching facilities have little influence on ideological and political learning ability. Therefore, the remaining nine factors and dependent variables that are significantly related to the ideological and political learning ability are selected into the classification variables of cluster analysis.

The sum of squares within the group represents the sum of squares of errors of the sample data and their mean values of each level or group, which reflects the dispersion of the observed values of each sample, also known as the sum of squares of errors. In this study, the sum of squares within each group is obtained by using Python language.

$$S_E = \sum_{p=1}^{r} \sum_{q=1}^{n_p} \left(y_{pq} - \bar{y}_p \right)^2, \tag{10}$$

where y_{pq} is a random variable within the group. y_p is the sample mean, and n is the sample size.

Cluster the above selected 10 factors, and calculate the sum of squares within the group corresponding to different cluster numbers, as shown in Figure 3. When the cluster number is 5, the sum of squares within the group basically does not change, so the cluster number k in this study is 5.

After 50 iterations and reclassification, the final clustering results are as follows.

3.5. Clustering Results and Related Suggestions. In order to improve the learning ability of postgraduates, this paper classifies postgraduates into 5 categories by cluster analysis based on the evaluation scores of each postgraduates' influencing factors on learning ability. Finally, according to the characteristics of each category, it puts forward specific suggestions for postgraduates to improve their learning ability. The results are as follows.

The first category: the theoretical knowledge of general industries. For this type of talents, we should start to create opportunities for practical application and encourage them to apply the theoretical knowledge they have learned to practice.

The second category: all-round knowledge-based high-end industries. The industry in which this type of talents is located has the highest requirements for ideological and political learning ability, and the difficulty and innovation degree of the research topics are high. For this class, the most important thing is to equip a strong team of teachers to improve their mathematical research ability.

The third category: other knowledge types in the literary industry. These individuals are at the lowest level in theoretical achievement and practical ability. For this kind of talents, cultivating interest in ideological and political learning is the easiest way to achieve.

The fourth category: the theoretical knowledge of middle-end industries. For this type of graduate students, schools should carry out more activities in ideological and political practice and improve ideological and political literacy to cultivate interest in learning.

The fifth category: other knowledge-based industries in high-end industries. This type of industry has a high demand for ideological and political education. For this type of graduate students, schools should also carry out more practical activities and lectures on improving literacy to cultivate interest and improve the atmosphere of learning.

4. Conclusion

The ideological and political theory course in colleges and universities is the main channel to strengthen and improve the ideological and political education of college students. This paper proposes an analysis algorithm based on BP neural network and improved k-means clustering, aiming to realize the research of students' innovative ability in ideological and political course. The algorithm collects the influencing factors of ideological and political learning ability. And BP neural network is used to quantify the learning ability and solve the problem of index weight. In the meantime, it avoids artificial subjective factors and ensures the reliability of evaluation results. Then, the correlation analysis screens out the influence of ideological and political learning ability on privacy. Finally, an improved DC k-means clustering algorithm is designed, which is applied to cluster-related ideological and political research individuals. The experimental results demonstrate that the improved clustering algorithm designed in this paper has high clustering accuracy. This paper can effectively analyze the learning ability of ideological and political course based on BP neural network and improved k-means clustering analysis algorithm. Finally, specific suggestions for each type of individual are listed out. The future research direction is to study the parameter model optimization method of BP neural network to further improve the accuracy of learning ability analysis.

Data Availability

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

Conflicts of Interest

The author declares no competing interests.

Acknowledgments

This work is supported by the Xijing University.

References

[1] G. Zhou, "Key problems and solutions of the application of artificial intelligence technology," in *Artificial Intelligence in China*, pp. 407–414, Springer, Singapore, 2020.

[2] J. Knox, "Artificial intelligence and education in China," *Learning, Media, and Technology*, vol. 45, no. 3, pp. 298–311, 2020.

[3] H. Du, "An English network teaching method supported by artificial intelligence technology and WBIETS system," *Scientific Programming*, vol. 2021, 9 pages, 2021.

[4] Z. Sun, M. Anbarasan, and K. D. Praveen, "Design of online intelligent English teaching platform based on artificial intelligence techniques," *Computational Intelligence*, vol. 37, no. 3, pp. 1166–1180, 2021.

[5] L. Huang, C. D. Wang, H. Y. Chao, J. H. Lai, and P. S. Yu, "A score prediction approach for optional course recommendation via cross-user-domain collaborative filtering," *IEEE Access*, vol. 7, pp. 19550–19563, 2019.

[6] Z. Chen, X. Liu, and L. Shang, "Improved course recommendation algorithm based on collaborative filtering," in *2020 International Conference on Big Data and Informatization Education (ICBDIE)*, pp. 466–469, IEEE, Zhangjiajie, China, 2020.

[7] R. Al-Shabandar, A. Hussain, A. Laws, R. Keight, J. Lunn, and N. Radi, "Machine learning approaches to predict learning outcomes in massive open online courses," in *2017 International Joint Conference on Neural Networks (IJCNN)*, pp. 713–720, IEEE, Anchorage, AK, USA, 2017.

[8] L. Chen, V. Jagota, and A. Kumar, "Research on optimization of scientific research performance management based on BP neural network," *International Journal of System Assurance Engineering and Management*, pp. 1–10, 2021.

[9] X. Ruan, Y. Zhu, J. Li, and Y. Cheng, "Predicting the citation counts of individual papers via a BP neural network," *Journal of Informetrics*, vol. 14, no. 3, p. 101039, 2020.

[10] S. Heidari, M. Alborzi, R. Radfar, M. A. Afsharkazemi, and A. Rajabzadeh Ghatari, "Big data clustering with varied density based on MapReduce," *Journal of Big Data*, vol. 6, no. 1, pp. 1–16, 2019.

[11] K. P. Sinaga and M. S. Yang, "Unsupervised K-means clustering algorithm," *IEEE Access*, vol. 8, pp. 80716–80727, 2020.

[12] J. García, V. Yepes, and J. V. Martí, "A hybrid k-means cuckoo search algorithm applied to the counterfort retaining walls problem," *Mathematics*, vol. 8, no. 4, pp. 555–577, 2020.

[13] A. Karlekar, A. Seal, O. Krejcar, and C. Gonzalo-Martin, "Fuzzy k-means using non-linear s-distance," *IEEE Access*, vol. 7, pp. 55121–55131, 2019.

[14] Z. Tang, Z. Zhu, Y. Yang, L. Caihong, and L. Lian, "D-K-means algorithm based on distance and density," *Application Research of Computers*, vol. 37, no. 6, pp. 1719–1723, 2020.

[15] S. Yu, X. Li, L. Zhao, and Q. Qiu, "Local density based potential dictionary construction for low rank representation in hyperspectral anomaly detection," in *Algorithms, Technologies, and Applications for Multispectral and Hyperspectral Imagery XXVI*, vol. 11392, p. 1139218, International Society for Optics and Photonics, 2020.

[16] J. Deng, Y. Deng, and K. H. Cheong, "Combining conflicting evidence based on Pearson correlation coefficient and weighted graph," *International Journal of Intelligent Systems*, vol. 36, no. 12, pp. 7443–7460, 2021.