Hindawi

Retraction

Retracted: Research on the Effectiveness of Labor Education for College Students Based on Big Data Technology

Advances in Multimedia

This article has been retracted by Hindawi following an investigation undertaken by the publisher [1].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] S. Yu, "Research on the Effectiveness of Labor Education for College Students Based on Big Data Technology," *Advances in Multimedia*, vol. 2022, Article ID 5328851, 11 pages, 2022.

Research Article

Research on the Effectiveness of Labor Education for College Students Based on Big Data Technology

Siyue Yu [ID]

School of Vehicle Engineering, Xi'an Aeronautical Institute, Xi'an, Shaanxi 710077, China

Correspondence should be addressed to Siyue Yu; 201509002@xaau.edu.cn

In order to analyze the effect of labor education (LE) for college students in college education, this paper combines big data technology to analyze the effectiveness of LE for college students. Moreover, this paper classifies the constraints of clustering effectiveness indicators and analyzes the characteristics and applications of various criteria in detail. In addition, this paper compares the effectiveness criteria of several commonly used fuzzy clustering based on relative criteria and compares and analyzes the performance and evaluation effect of these criteria. On this basis, this paper proposes an improved effectiveness criterion. Through the controlled experiment, it can be seen that the LE for college students can effectively improve the various abilities of college students, and it has a certain auxiliary effect on the learning and growth of college students.

1. Introduction

Social practice, voluntary service, scientific and technological activities, production practice, and other forms enrich and expand the scope of students' participation in labor. However, due to the influence of living environment, family education, social atmosphere, and other aspects, some students are drifting away from labor, and they are in the embarrassing situation of "it's important to say, but not to do." It is mainly manifested in the following aspects. First of all, they lack the correct understanding, that is, they "look down on labor," despise labor, especially manual labor, and think that participating in labor is inferior. Under the cover of "study," some students only participate in the required internships and practices, and stay away from other labor. Second, they lack labor awareness, that is, "don't want to work." Compared with their parents, the material living conditions have been greatly improved, so that many students lack intuitive feeling and personal experience of "work creates wealth" and "work creates a better life." Not only do they not take the initiative to participate in family labor, school labor, and social labor, but even after graduation, they still want to get something for nothing, and the phenomenon of "gnawing on old age" appears, which is undoubtedly closely related to the dominance of exam-oriented education and the absence of LE [1]. Finally, they lack labor skills, that is, "cannot work," and they usually participate in less labor and lack labor skills training. Throughout a person's educational experience, LE shows a trend of gradually weakening from low to high. Therefore, as a higher education stage directly connected to the society, the labor, especially manual labor, that students participate in is often the least [2].

Labor mainly includes daily labor, production labor, and service labor. From the perspective of the connotation of LE, literature [3] believes that LE is an education combined with the labor process, so that students can master labor knowledge skills and labor scientific knowledge and improve their labor literacy. Literature [4] pointed out that LE includes the cultivation of labor values and spirit, and the cultivation of labor knowledge, skills, and abilities. Literature [5] believes that LE is to combine life, practice and education, pay attention to life, and pay attention to the cultivation of vitality. Literature [6] believes that LE aims to educate and cultivate people through labor. The role of LE is to educate the majority of students to recognize and strive to practice the spirit of labor, model workers, and craftsmen. Literature [7] pointed out that "LE should focus on education and cultivate students' feelings for the working people. If they only study

but do not work, do not contact social practice, and do not understand the creation of social wealth, it is not conducive to their healthy growth and all-round development." LE has a fundamental and overall status in the construction of the education system and plays a fundamental role in the training of socialist builders and successors. Literature [8] pointed out that the combination of education and labor is an effective way to realize the all-round development of human beings, and integrating LE into the whole process of talent training in colleges is conducive to the all-round development of students. LE is an important part of the national human capital strategy in the new era, and it is an inevitable requirement for cultivating students with all-round development of morality, intelligence, physique, beauty, and labor. LE is an important part of the comprehensive education system, which is conducive to laying the foundation for the four educations of morality, intelligence, physical fitness, and beauty. Literature [9] believes that LE helps to solve the problem of the disconnection between higher education personnel training and social needs.

Existing studies have pointed out that there are deficiencies in LE for students in terms of content, methods, and approaches, and the effect of LE is not good. Some colleges limit the scope of LE services to the campus and are limited to labor behaviors such as environmental cleaning and maintenance, duty, and assistance with teaching affairs. The educational content is not rich enough. Wrong ideological cognitions such as "lower class" and "opposition between labor and study" [10]. The lack of understanding of the connotation of LE, the inaccurate positioning of practice, and the lack of campus atmosphere have led to the poor effect of LE in colleges, the lack of labor skills, weak labor awareness, and utilitarian labor concepts of students [11]. Literature [12] pointed out that the status of LE of teaching system of colleges is not high, which is reflected in the lag in the cultivation of LE skills, the loss of the spiritual value of LE, and the shortage of LE curriculum resources. Literature [13] believes that LE in colleges has the phenomenon of dwarfing the value of LE, virtual LE mechanism, and narrowing of LE content. The problem of fuzziness, weakening, narrowness, and dilution of physical education in colleges and universities reflects the value dilemma of "inherent" physical education concepts, the "Deviation" of educational objectives, the "one-way" of educational methods, and the insufficient implementation of LE experience and deep participation.

Labor is both a means and an end of education. LE is an important part of talent training in colleges, and it is the organic unity of labor knowledge education and labor process education [13]. LE has the basic characteristics of being comprehensive, contemporary, and practical. LE should be incorporated into the talent training system, integrated into the teaching of professional courses, connected to innovation and entrepreneurship education, and connected to the second classroom for students [14]. Literature [15] pointed out that the key to promoting LE lies in strengthening ideological guidance to highlight the orientation of LE, mobilizing students' enthusiasm to strengthen the recognition of LE, and joint management by all parties to strengthen the synergy of LE, expand the content of education to enhance the recognition of LE, and strengthen the affinity

of LE. Literature [16] believes that the development of LE should strengthen the top-level design of LE in colleges and improve the practice system of LE in colleges. Promote the standardization, normalization, and sustainable development of LE from four aspects: colleges, society, families, and students themselves. LE in colleges should follow the internal logic of cognitive identification, value identification, and behavior identification and exert its talent training function by strengthening value guidance, creating a good atmosphere, returning to the life world, and strengthening practical experience [17]. Literature [18] believes that in carrying out LE, it is necessary to clarify the scientific connotation of LE in colleges in terms of ideological understanding, education subject, and educational content; strengthen the top-level design of LE in colleges; make classrooms the main channel of LE; expand LE platforms; and improve the mechanism to encourage other organizations in the school to actively participate. Literature [19] pointed out that LE should be implemented through mechanisms such as professional study and social practice, moral literacy and daily practice, tempering of character and hard training, and the combination of entrepreneurial employment and value realization, so as to guide students to establish correct labor values and improve professional skills, cultivate labor emotion, and cultivate labor morality.

This paper combines the big data technology to analyze the effectiveness of LE for students, improve the effect of LE for students, and promote the stable development of college education.

2. Effective Clustering Analysis Based on Big Data Technology

2.1. Clustering Effectiveness Criterion Classification and Analysis. Generally speaking, clustering validity evaluation methods. Among them, both the outer criterion and the inner criterion are based on statistical tests, and both have high computational complexity. The purpose of the criteria in both approaches is to measure how well a dataset matches a given prior structure. The relative evaluation rule is to evaluate the clustering results; it seeks the parameter values and clustering patterns that can make the clustering algorithm get the best clustering results. In addition, if it is divided according to the object of clustering effectiveness criterion evaluation.

The external evaluation only pays attention to the degree of agreement between the clustered partition and the known partition. Furthermore, external evaluation requires that the structure and classification of the clustered dataset be known. Although this method is relatively simple and straightforward, it ignores the expected characteristics of the clustering results, so such criteria are very straightforward in evaluation. In the following, several commonly used external evaluation criteria are introduced.

2.1.1. F-Measure Criteria. This criterion involves two concepts in information retrieval: precision and recall. We assume that the dataset is j and each category is i. Moreover, it is known that the number N_j of all data objects in dataset j and the number N_i, N_{ij} of data objects in a certain category i represent the actual number of categories of category i in

dataset j after clustering. Then, the precision and recall of the clustering results of category i in dataset j are defined as follows:

$$p = \text{precision}(i, j) = \frac{N_{ij}}{N_i},$$

$$r = \text{recall}(i, j) = \frac{N_{ij}}{N_j}, \tag{1}$$

$$F(i) = \frac{2pr}{(p + r)}.$$

In terms of clustering results, F-measure is used as the evaluation value of category i, so which clustering result is used as the mapping of category i is determined by the value of F-measure. According to the definition, when a certain clustering result obtains the highest F-measure value, the cluster is regarded as the mapping of classification i. The weighted average of the F-measure of each category i can obtain the total F-measure evaluation value of a certain clustering result, which is defined as follows:

$$F = \frac{\sum_i [|i| \times F(i)]}{\sum_i |i|}. \tag{2}$$

Among them, $|i|$ is the number of all objects in category i. The value of F ranges from 0 to 1, and the higher the value, the better the clustering effect.

2.1.2. Rand Criteria. It is defined as follows: if the predicted structure of the dataset X is divided into $X = \{X_1, X_2, \cdots, X_s\}$, and the clustering result is $C = \{C_1, C_2, \cdots, C_c\}$, the quality of the clustering result is evaluated by comparing X and C. First, the following items are defined. Among them, SS and DD are used to evaluate the identity of the two clusters, and SD and DS are used to evaluate the difference of the two clusters. Rand criteria can be expressed as follows:

$$\text{Rand} = \frac{SS + DD}{SS + SD + DS + DD}. \tag{3}$$

In further calculations, the similarity between X and C can be represented by the Jacard coefficient:

$$J = \frac{SS}{SS + SD + DS}. \tag{4}$$

The value range of the above two criteria is [0, 1]. The size of the Rand criteria value directly reflects the degree of agreement between the clustering results and the known divisions.

The unknown structure of the dataset makes the internal evaluation of the clustering results not as direct as the external evaluation, which can only be evaluated from the attributes of the clustering results, namely, the intraclass identity and the interclass separation. Moreover, the size of the clusters often has a certain influence on the clustering results, so it also needs to be considered in order to balance the clustering results. Several common internal evaluation methods are introduced.

(1) Intracluster variance

This criterion mainly focuses on obtaining the best clustering effect through the minimum distance within the cluster and uses the idea of the minimum sum of squares of errors to find the minimum distance within the cluster. C is all clusters of the dataset, v_k is the center of a cluster c_k, and $d(x, v)$ represents the distance between objects; then, the intracluster distance of the data object x is defined as follows:

$$V(C) = \sum_{C_k \subset C} \sum_{x \in C_k} d(x, v_k)^2. \tag{5}$$

In order to achieve the best clustering effect, the variance value within the cluster should be as small as possible, and the clustering division should be better at this time.

(2) Hubert's Γ statistics

If X is defined as the dataset matrix, the adjacent matrix Y is defined as follows:

$$Y(i, j) = \begin{cases} 1 & \text{When } x_i \text{ and } y_i \text{ belong to different clusters } i, j = 1, 2, \cdots, N, \\ 0 & \text{other.} \end{cases} \tag{6}$$

The evaluation of the validity of the criteria for the clustering results is mainly determined by comparing the degree of consistency between the clustering results and the adjacent matrix Y; then,

$$\Gamma = \left(\frac{1}{M}\right) \sum_{i=1}^{N-1} \sum_{j=i+1}^{N} X(i, j) Y(i, j). \tag{7}$$

The similarity between X and Y increases as the value of Γ increases.

Compared with the above two evaluation methods, relative evaluation no longer simply focuses on the structure of the dataset. On the basis of determining the clustering algorithm, the clustering results are evaluated by changing the parameter values of the clustering algorithm, and finally, the purpose of finding the optimal parameter setting and clustering mode is realized. The following describes the critical evaluation criteria:

(1) SD validity. It is a relative evaluation method that uses the concepts of the overall scatter of clusters and the degree of separation between classes to linearly combine the distances within and between classes

If the variance of the dataset X is assumed to be $\sigma(x)$, its pth dimension variance is defined as follows:

$$\sigma_x^p = \frac{1}{n \sum_{k=1}^{n} \left(x_k^p - \overline{x^p}\right)^2},$$

$$\sigma^p(v_i) = \frac{1}{n_i \sum_{k=1}^{n} \left(x_k^p - v_i^p\right)^2}. \tag{8}$$

From the above content, it can be seen that the average scatter of clusters is defined as follows:

$$\text{Scat}(c) = \frac{1}{c\sum_{i=1}^{c}\|\sigma(v_i)\|/\|\sigma(X)\|},$$

$$\text{Dis}(c) = \frac{D_{\max}}{D_{\min}\sum_{k=1}^{c}\left(\sum_{i=1}^{c}\|v_k - v_i\|\right)^{-1}}. \tag{9}$$

Finally, SD criteria can be defined as follows:

$$\text{SD}(c) = a\text{Scat}(c) + \text{Dis}(c). \tag{10}$$

Among them, $\alpha = \text{Dis}(c_{\max})$ is the weighting factor, and the maximum number of clusters that need to be input is $c_{\max}$.

(2) Dunn criteria

Different from SD criteria, Dunn criteria does not use a linear combination method, but a nonlinear combination method using the ratio of intraclass diameter and interclass distance to evaluate the clustering results. $d(v_i, v_j) = \min_{x\in v_i, y\in v_j} d(xy)$ represents the dissimilarity between class v_i and class v_j. If the diameter $\text{diam}(v) = \max_{x,y\in v} d(x,y)$ is the maximum intraclass distance, then

$$D(v) = \min_{i=1,,\cdots,c}\left(\min_{j=i+1,\cdots,c}\left(\frac{d(v_i, v_j)}{\max_{k=1,2,\cdots,c}\text{diam}(v_k)}\right)\right). \tag{11}$$

The interclass distance is large; the expected effect is better. That is, the larger the Dunn value, the better the clustering effect. However, since this criterion mainly measures distance, outliers have a greater impact on the criterion.

2.2. Several Common Fuzzy Clustering Validity Criteria. Due to the increasing complexity of datasets, it is becoming more and more difficult to clearly divide the data in the datasets. Therefore, due to the limitation of its algorithm idea, the hard clustering algorithm can no longer meet the current dataset division requirements. In response to this situation, the idea of fuzzy partitioning came into being, and the fuzzy clustering algorithm was also widely used. To evaluate the effectiveness of fuzzy clustering algorithms, fuzzy evaluation criteria are proposed. This paper focuses on the fuzzy evaluation criteria and its application and firstly introduces several commonly used fuzzy evaluation criteria.

The partition coefficient (PC) is the first cluster validity criterion designed with good mathematical properties and practicality, which is used to measure the degree of overlap between clusters. u_{ij} is the membership degree of data x_j in cluster i, the definition is as follows:

$$V_{\text{pc}} = \frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^2. \tag{12}$$

It can be seen from the definition that with the increase of the number of cluster centers c, the value of V_{pc} will show a monotonic trend. Although c is considered to obtain the optimal number of clusters when V_{pc} is the largest, because of the monotonic variation of V_{pc}, it often does not provide a good reference to use V_{pc} to determine the optimal number of clusters.

Partitioning entropy (PE) is also an early proposed validity criterion. It defines the number of data n, the number of clusters c, the degree of membership matrix U, and the degree of membership u_{ij} of the data x_j in cluster i.

$$V_{\text{PE}} = -\frac{1}{n}\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}\log u_{ij}. \tag{13}$$

Contrary to V_{pc}, the optimal partition can be obtained when V_{PE} takes the minimum value. However, as can be seen from the definition, V_{PE} as the criterion for determining the optimal number of clusters is also increasingly limited as the complexity of the data increases.

The dataset is $X = \{x_j/j = 1, 2, \cdots, n\}$, v_i, where $i = 1, 2, \cdots, c$ is the center of the ith class and u_{ij} is the fuzzy membership of the data x_j to class i.

The Xie-Beni criteria are as follows:

$$V_{\text{xie}}(c) = \frac{\pi}{s} = \frac{1/n\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \|x_j - v_i\|^2}{\min_{1\leq p,q\leq c, p\neq q}\|v_p - v_q\|^2}. \tag{14}$$

From the definition, it can be found that $V_{\text{xie}}(c)$ considers both the intraclass compactness and the interclass separation of clusters. Among them, $\min_{1\leq p,q\leq c, p\neq q}\|v_p - v_q\|^2$ is used to evaluate the degree of separation between classes, the larger the value, the better the clustering effect. $1/n\sum_{i=1}^{c}\sum_{j=1}^{n} u_{ij}^m \|x_j - v_i\|^2$ is used to evaluate intraclass tightness. The smaller the value, the better the clustering effect. Therefore, the optimal clustering is obtained when $V_{\text{xie}}(c)$ takes the minimum value. From this, it can be obtained that when c^* satisfies $V_{xie}(c^\circ) = \min_c(V_{xie}(c))\}$, the value of c^* is the best at this time.

In the following, we specifically dissect this formula to see how it reflects this problem.

If $d_{ij} = u_{ij}\|x_j - v_i\|$ is used to represent the fuzzy weighted membership degree of the distance from the data x_j to the center of the ith class, that is, the fuzzy deviation, then the number of fuzzy data in the ith class is expressed as follows:

$$n_i = \sum_{j=1}^{n} u_{ij},$$

$$\sum_{i=1}^{c} n_i = n. \tag{15}$$

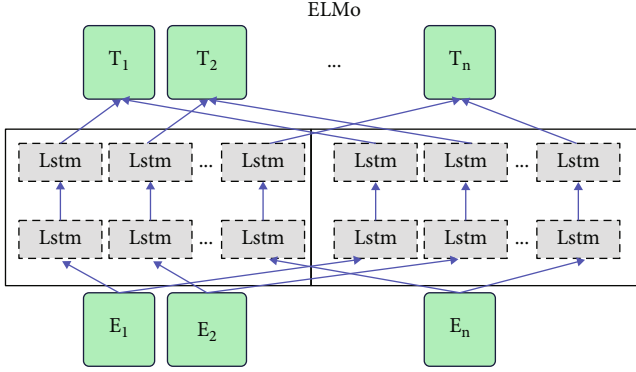Among them, the variation of the ith class is expressed as follows:

ELMo

Figure 1: Structure diagram of the ELMo model.

$$\sigma_i = \sum_{j=1}^{n} d_{ij}^2. \tag{16}$$

The mean variation of the ith class, the tightness, is expressed as follows:

$$\pi_i = \frac{\sigma_i}{n_i}. \tag{17}$$

From this, the total variation of dataset X can be expressed as the sum of the variation of each class:

$$\sigma = \sum_{i=1}^{c} \sigma_i. \tag{18}$$

It can be seen from the formula that the total variation of the dataset X at this time is the same as the objective function when the fuzzy factor $m = 2$ in the FCM algorithm.

The tightness of the fuzzy C partition of dataset X is

$$\pi = \frac{\sigma}{n}. \tag{19}$$

π reflects the overall distribution characteristics of the dataset structure. The size of the value determines the tightness of the cluster. The smaller the value, the higher the overall tightness of the cluster.

The degree of separation of fuzzy C partitions is expressed by the minimum distance between each class center:

$$s = \min_{1 \le i, j \le c, i \ne j} \left\| v_i - v_j \right\|^2. \tag{20}$$

The $V_{\text{kwon}(c)}$ validity criterion is defined as follows:

$$V_{\text{kwon}}(c) = \frac{\sum_{i=1}^{c} \sum_{j=1}^{n} u_{ij} \left\| v_j - x_i \right\|^2 + (1/c) \sum_{i=1}^{c} \left\| v_i - \bar{v} \right\|^2}{\min_{1 \le i, j \le c, i \ne j} \left\| v_i - v_j \right\|^2}. \tag{21}$$

Among them, $\bar{v} = 1/n \sum_{j=1}^{n} X_j$.

For the Xie-Beni criteria, when the number of clusters c gradually increases and approaches the total number of data in the dataset, there is a decreasing trend, which leads to the defect of losing the evaluation ability. The Xie-Beni criteria is improved by introducing a penalty function $1/c \sum i \left\| v_i - v \right\|^2$. After the adjustment of the penalty function, the criterion can suppress the decrease of the criterion value to a certain extent. As with the Xie-Beni criteria, when the criterion value is the smallest, the clustering result is optimal.

(1) Rezaee criteria

Vcwb (Compose within and between scattering) is defined as follows:

$$\text{Vcwb}(c) = \alpha \, \text{Scat}(c) + \text{Dis}(c). \tag{22}$$

Among them, $\alpha = \text{Dis}(c_{\max})$.

$$\text{Scat}(c) = \frac{1/c \sum_{i=1}^{c} \left\| \sigma(v_i) \right\|}{\left\| \sigma(X) \right\|}. \tag{23}$$

Among them,

$$\begin{aligned}
\|X\| &= (X^\tau X)^{1/2}, \\
\bar{x} &= \frac{1}{n} \sum_{k=1}^{n} x_k, \\
\sigma(X) &= \frac{1}{n} \sum_{k=1}^{n} (x_k - \bar{x})^2, \\
\sigma(v_i) &= \frac{1}{n} \sum_{k=1}^{n} u_{ki}(x_k - v_i)^2.
\end{aligned} \tag{24}$$

The degree of separation is

$$\text{Dis}(c) = \frac{D_{\max}^2}{D_{\min}^2} \sum_{i=1}^{c} \left(\sum_{j=1}^{c} \left\| v_i - v_j \right\|^2 \right)^{-1}. \tag{25}$$

Among them, $D_{\min} = \min_{i \ne j} \left\| v_i - v_j \right\|$, $D_{\max} = \max_{i \ne j} \left\| v_i - v_j \right\|$, $i, j \in [1, c]$.

Rezaee makes up for the shortcomings of the traditional criterion function in terms of "closeness" and "separation" to a certain extent by combining the linear combination method with the scaling factor scaling idea. However, due to the limited processing power of this scaling, the huge gap in the values of "closeness" and "separation" has not been fundamentally resolved.

2.3. An Improved Clustering Effectiveness Criteria (CS). Aiming at the deficiencies of several commonly used criteria, this paper improves the validity criterion on the basis of XB criteria and proposes an improved validity criterion by combining the ideas of "closeness" and "separation." The criterion can avoid the monotonous change of the criterion with the increase of the number of cluster, and can effectively make up for the numerical gap between the compactness and the separation degree.
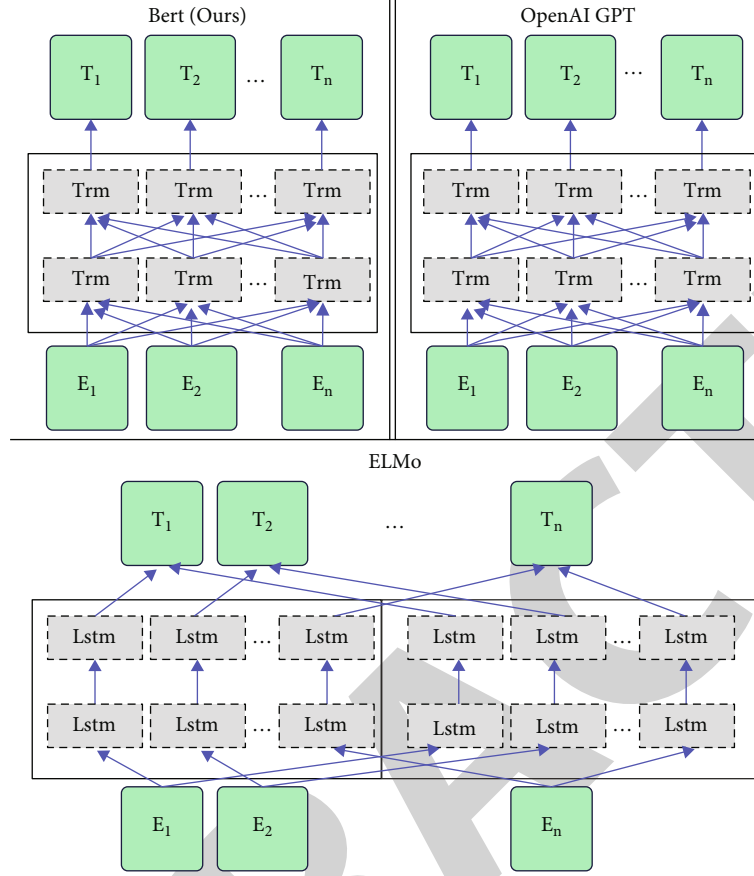
FIGURE 2: Structure comparison diagram of the three models.

In fuzzy division, $X = \{x_1, x_2, x_3, \cdots, x_n\}$ is the n-element dataset.

$$C(c, U) = \sum_{i=1}^{c} \frac{\sum_{j=1}^{n} (u_{ij})^m \|x_j - v_i\|}{n_i} * \left(\frac{c+1}{c-1}\right)^{1/2},$$

$$\lim_{c \longrightarrow n} \|x_j - v_i\|^2 = 0. \tag{26}$$

Among them, $\sum_{j=1}^{n} (u_{ij})^m \|x_j - v_i\|$ is the intraclass sum of squared errors based on the Euclidean distance. As c increases, the distance from each data to the center of the class decreases. This makes $1/n_i = 1/\sum_{j=1}^{N} u_{ij}$, as the weight of each cluster, show an increasing trend, which can limit the monotonically decreasing situation of the tightness measure. $(c + 1/c - 1)^{1/2}$ will decrease with the increase of c, and it cooperates with m to achieve the best clustering effect.

In fuzzy clustering, in order to achieve better clustering effect, we should make the overlapping sample data as little as possible, and the degree of separation between classes should be as high as possible. Therefore, the degree of separation between classes can also be measured by the degree of overlap between classes.

S_{c1} and S_{c2} are two fuzzy clusters belonging to the fuzzy partition (U and V). The membership degrees of sample x_i in S_{c1} and S_{c2} are denoted as $S_{c1}(x_i), S_{c2}(x_i)$, respectively. It can be seen that the overlap degree of sample x_i in S_{c1} and S_{c2} is

$$S(S_{c1}, S_{c2} : x_i) = \min (S_{c1}(x_i), S_{c2}(x_i)). \tag{27}$$

The total overlap degree of fuzzy cluster S_{c1} and S_{c2} is

$$S(S_{c1}, S_{c2}) = \sum_{k=1}^{n} S(S_{c1}, S_{c2} : x_k) \times \omega(x_k). \tag{28}$$

Among them, $\omega(x_i) = -\sum_{i=1}^{c} u_{s_k}(x_k) \log_a u_{s_k}(x_k)$, and this weight index is mainly used to adjust the overlapping part of the data points. Also, balancing the effect of some clusters with high or low overlap on the overall overlap. Then, the degree of cluster separation is defined as the average degree of overlap. Since every two clusters may overlap, the formula for the average degree of overlap is defined as follows:

$$S_{\text{heavy}}(c, U) = \frac{2\sum_{i \neq j}^{c} S(S_{ci}, S_{cj})}{nc(c-1)},$$

$$S(c, U) = 1 - S_{\text{heavy}}(c, U). \tag{29}$$

It can be seen from the formula that the lower the degree of overlap in the clustering, the higher the degree of separation between clusters. At this time, $S_{\text{heavy}}(c, U)$ reaches the
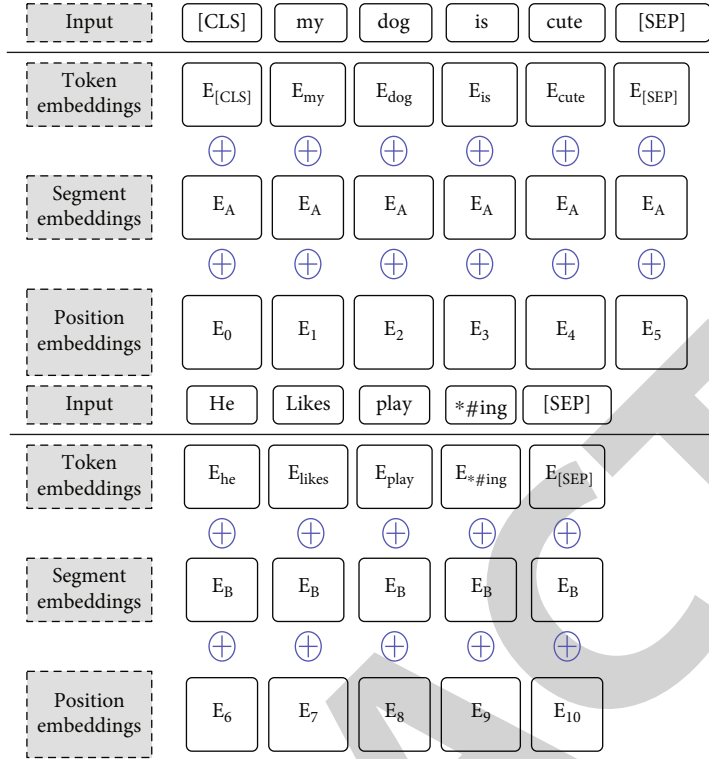
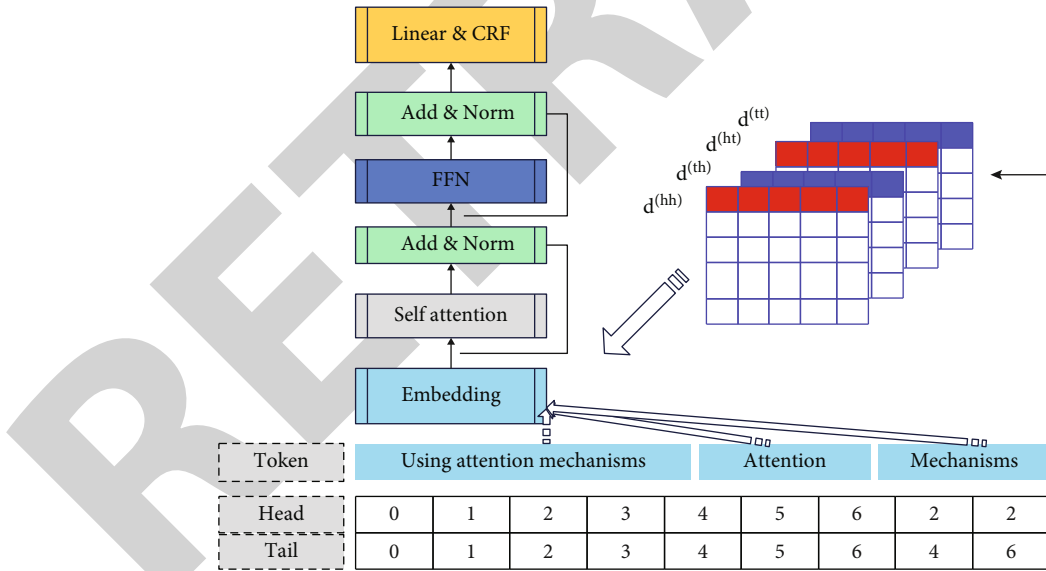FIGURE 3: Composition of input vectors of the Bert model.

FIGURE 4: The model structure diagram of the imported dictionary.

minimum and the value of $S(c, U)$ reaches the maximum. Therefore, when the degree of separation is higher, it is larger, and the clustering effect is better. At this time, the optimal number of clusters is obtained.

Since the measurement scalars of separation degree and closeness are different, in order to carry out calculation and comparison, it is necessary to normalize the two. The result can be expressed as follows:

$$
\begin{aligned}
\text{Com}(c, U) &= \frac{C(c, U)}{\max\left\{C(c_1, U), \cdots, C(c_{\max}, U)\right\}}, \\
\text{Sep}(c, U) &= \frac{S(c, U)}{\max\left\{S(c_1, U), \cdots, S(c_{\max}, U)\right\}}.
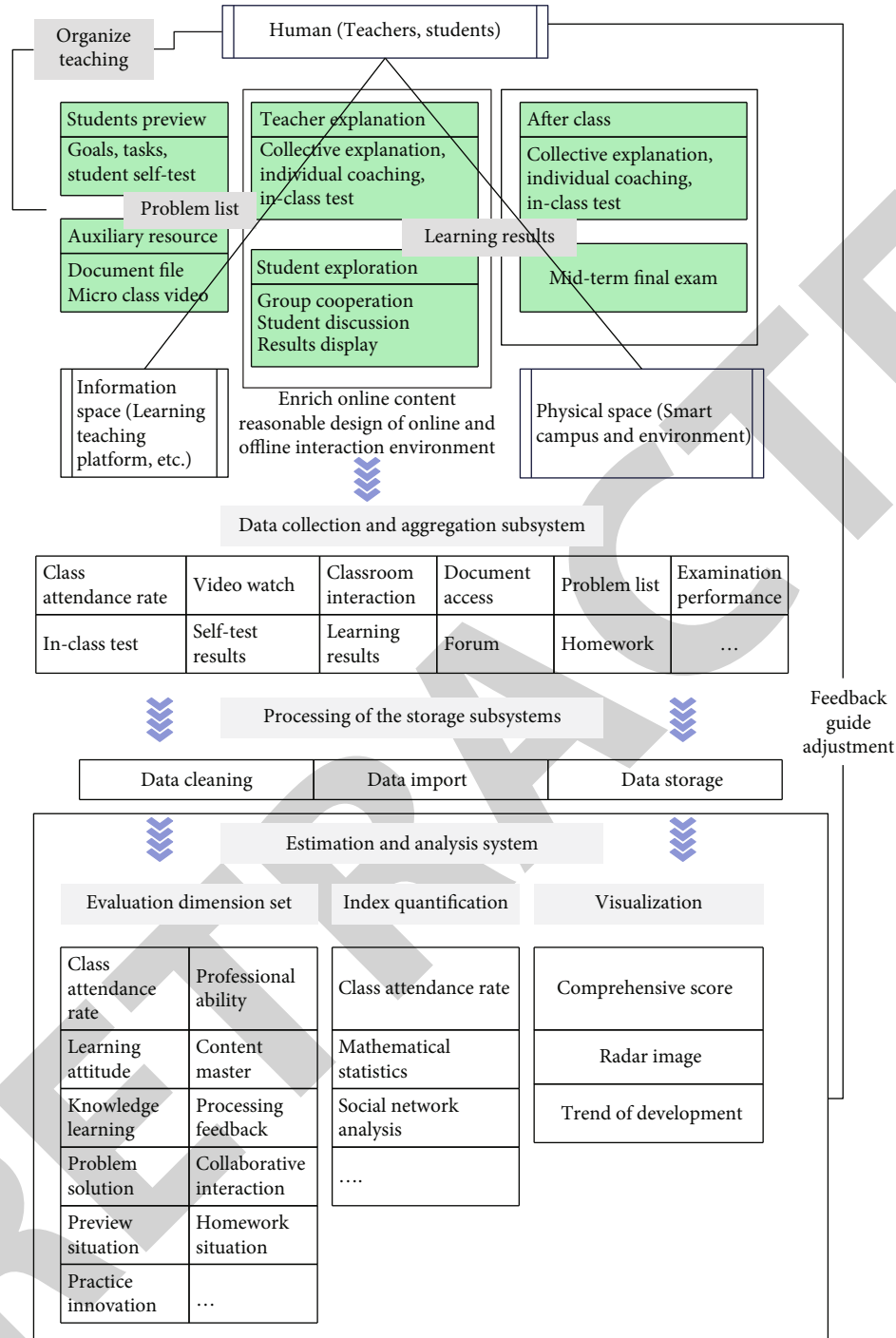\end{aligned}
\tag{30}
$$

FIGURE 5: Design model of the LE evaluation system of students.

The validity criterion CS can be expressed as follows:

$$CS = \frac{Com(c, U)}{Sep(c, U)}. \tag{31}$$

To sum up, in order to optimize the criterion CS, that is, to minimize the $Com(c, U)$ value. At the same time, $Sep(c, U)$ reaches a smaller degree of overlap, that is, the value of $Sep(c, U)$ reaches a larger value, and at this time, CS obtains a smaller value. Therefore, the optimal number of clusters is the number of clusters corresponding to the minimum value of CS.

3. Analysis of the Effectiveness of LE for Students Based on Big Data Technology

This paper combines the intelligent model to analyze the effectiveness of LE for students. After completing the training of this college student LE model, for each target
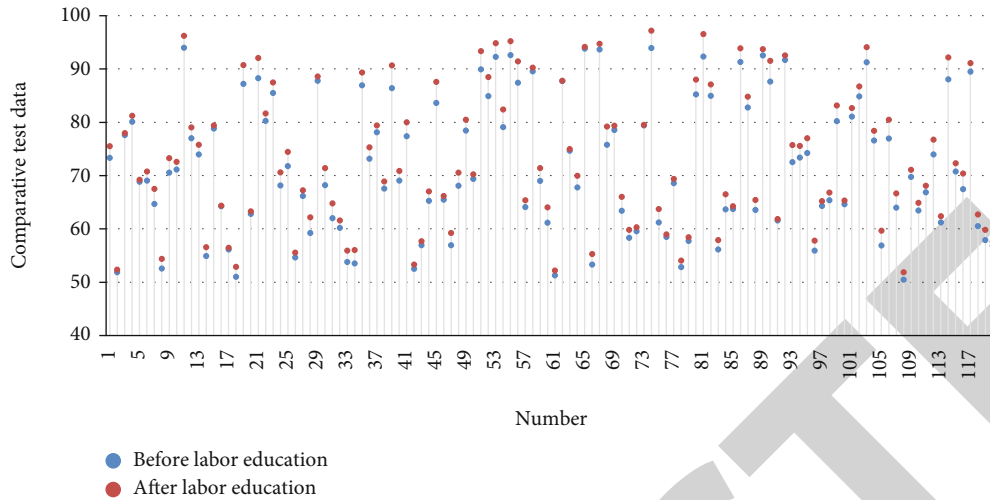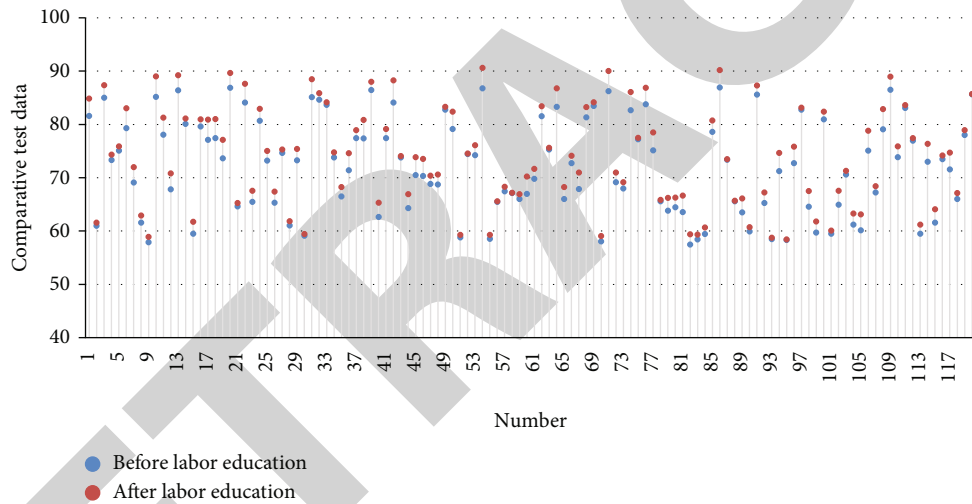
Figure 6: Comparison of learning enthusiasm.

Figure 7: Comparison of physical fitness.

vocabulary, the ELMo model calculates each intermediate layer of the bidirectional model and sums them up as a vector representation of the vocabulary. The structure of the ELMo model is shown in Figure 1.

The difference between a pretrained language model like ELMo and other previous language models is the traditional word vector, and a word corresponds to only one word vector. However, the ELMo model uses the bidirectional LSTM model framework, which can obtain the word vector representation that comprehensively considers the context from the model according to the specific input sentence, so as to obtain better representation ability. While the ELMo model is groundbreaking, it still needs some effort from practical application. This paper selects the data processing model by comparing multiple models, and Figure 2 shows the comparison of the three models.

A large part of the reason why the Bert model can be so successful is that it uses the Transformer encoder. The core

of the Transformer encoder is a fully connected self-attention mechanism. This mechanism can not only obtain the encoded representation of each word in the current context but also achieve sufficient interaction between each pair of words in the sentence, enhancing its expressive ability. In order to solve the problem that the relative position importance is seriously weakened, when each two words interact within the self-attention mechanism, a new input encoding method is proposed. The input vector composition of the Bert model is shown in Figure 3:

The final input consists of three parts: the first part is the normal word vector encoding, using the WordPiece word vector; the second part is the position vector encoding, which trains a position vector separately for each position; the third part is the sentence segmentation vector, which is used to deal with some tasks with different input forms.

Since the Bert model does not support adding dictionaries directly, in order to introduce course dictionaries, we change
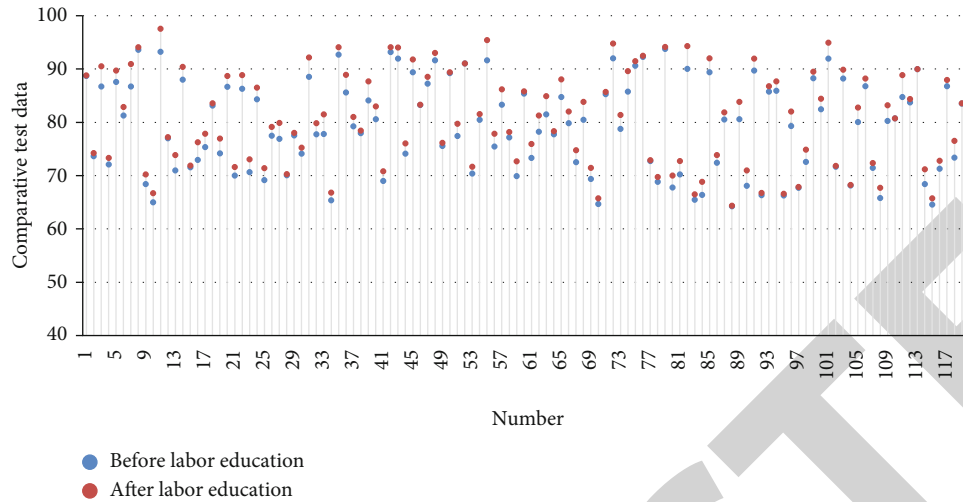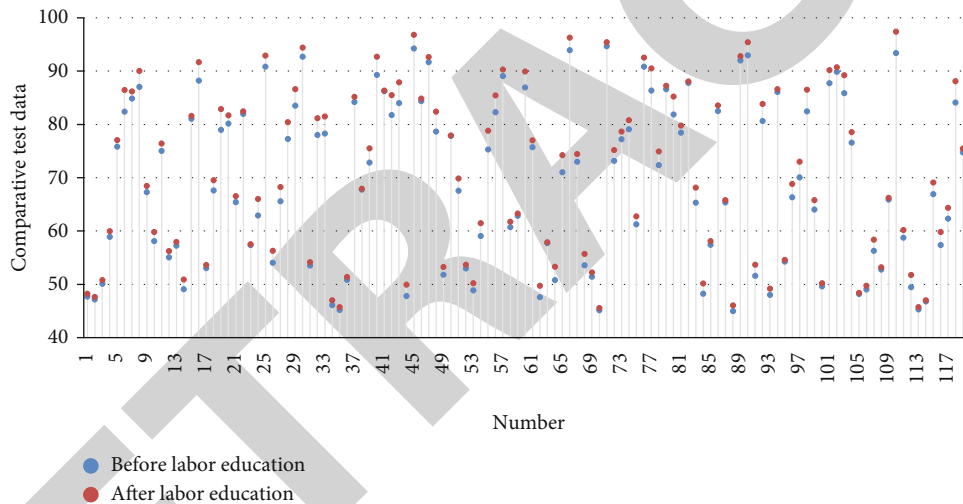
Figure 8: Comparison of compressive capacity.

Figure 9: Comparison of the degree of learning effort.

the original model and added a module for importing dictionary query results. The model structure is shown in Figure 4.

On the basis of the learning evaluation model and evaluation criteria analysis, according to the information system development theory, a learning evaluation system model in a smart campus environment is constructed. The system consists of a process data acquisition subsystem, a learning evaluation subsystem, and an evaluation result visualization subsystem, as shown in Figure 5. Under these subsystems, process data has undergone collection, aggregation, processing, evaluation, and visualization, and transformed from basic raw data into evaluation results that can be intuitively understood by students. It allows students to understand the personal learning process, allows teachers to better understand the learning status of students, and facilitates the teaching management department to monitor the quality of teaching and complete teaching management. The design model of the LE evaluation system for students is shown in Figure 5.

On the basis of the above, the evaluation of the LE system of students is carried out through multiple sets of data. Through the comparative test method, the students' enthusiasm for learning, physical quality, ability to resist pressure, and the degree of learning effort before and after LE are compared, and the results are obtained as shown in Figures 6–9 below.

Through the controlled experiment, it can be seen that the LE for students can effectively improve the various abilities of students, and it has a certain auxiliary effect on the learning and growth of students.

4. Conclusion

The LE for students in the new era has distinct characteristics of the times, and it is necessary to innovate the practice path of LE for students. The key to the construction of LE curriculum in the new era lies in the construction purpose, content, and

method. It needs to move towards cross-border integration, and it is necessary to explore the promotion path of LE in the new era. Moreover, colleges should design highly targeted LE courses according to different professional characteristics. In addition, the implementation of LE needs to be guaranteed by relevant LE policies. This paper combines the big data technology to analyze the effectiveness of LE for students to improve the effect of LE for students. Through the controlled experiment, it can be seen that the LE for students can effectively improve the various abilities of students, and it has a certain auxiliary effect on the learning and growth of students.

Data Availability

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

Conflicts of Interest

The author declares no competing interests.

Acknowledgments

References

[1] Y. Sun and Z. Liu, "The value of labor education in promoting the all-round development of college students and its implementation path," *International Journal of Social Science and Education Research*, vol. 3, no. 4, pp. 60–63, 2020.

[2] D. Li, "Investigation and analysis on the effect of labor education in Wenzhou higher vocational colleges in the new era," *International Journal of Social Science and Education Research*, vol. 4, no. 3, pp. 92–101, 2021.

[3] X. Zhou and K. Liu, "Research on the education mechanism of integration of labor education and ideological and political education in colleges in the new era," *International Journal of Social Science and Education Research*, vol. 5, no. 3, pp. 556–560, 2022.

[4] Y. Lu and H. Song, "The effect of educational technology on college students' labor market performance," *Journal of Population Economics*, vol. 33, no. 3, pp. 1101–1126, 2020.

[5] D. Xu, S. S. Jaggars, J. Fletcher, and J. E. Fink, "Are community college transfer students "a good bet" for 4-year admissions? Comparing academic and labor-market outcomes between transfer and native 4-year college students," *The Journal of Higher Education*, vol. 89, no. 4, pp. 478–502, 2018.

[6] M. Reyes, A. Dache-Gerbino, C. Rios-Aguilar, M. Gonzalez-Canche, and R. Deil-Amen, "The "geography of opportunity" in community colleges: the role of the local labor market in students' decisions to persist and succeed," *Community College Review*, vol. 47, no. 1, pp. 31–52, 2019.

[7] M. S. Giani, P. Attewell, and D. Walling, "The value of an incomplete degree: heterogeneity in the labor market benefits of college non-completion," *The Journal of Higher Education*, vol. 91, no. 4, pp. 514–539, 2020.

[8] G. Wiggan, D. Smith, and M. J. Watson-Vandiver, "The national teacher shortage, urban education and the cognitive sociology of labor," *The Urban Review*, vol. 53, no. 1, pp. 43–75, 2021.

[9] J. D. Collins, A. Manning-Ouellette, B. Neal, and M. Daglaris, "Our labor of love: socially just leadership education as civic leadership development," *New Directions for Higher Education*, vol. 2021, no. 195-196, pp. 153–162, 2021.

[10] O. Gurantz, "Impacts of state aid for nontraditional students on educational and labor market outcomes," *Journal of Human Resources*, vol. 57, no. 1, pp. 241–271, 2022.

[11] H. Wang, "Research on thinking obstacles from the perspective of labor economics," *Psychiatria Danubina*, vol. 33, suppl 8, pp. 278-279, 2021.

[12] L. D. Gonzales and D. F. Ayers, "The convergence of institutional logics on the community college sector and the normalization of emotional labor: a new theoretical approach for considering the community college faculty labor expectations," *The Review of Higher Education*, vol. 41, no. 3, pp. 455–478, 2018.

[13] C. Linder, S. J. Quaye, A. C. Lange, R. E. Roberts, M. C. Lacy, and W. K. Okello, "" A student should have the privilege of just being a student": student activism as labor," *The Review of Higher Education*, vol. 42, no. 5, pp. 37–62, 2019.

[14] J. Choi and H. Bae, "Changes in early labor market outcomes among young college graduates in South Korea," *The Annals of the American Academy of Political and Social Science*, vol. 688, no. 1, pp. 115–136, 2020.

[15] D. Xu and F. X. Ran, "The impacts of different types of college instructors on students' academic and labor market outcomes," *Journal of Policy Analysis and Management*, vol. 40, no. 1, pp. 225–257, 2021.

[16] C. Maggio and P. Attewell, "The surprising labor market success of part-time community college students," *Community College Journal of Research and Practice*, vol. 44, no. 7, pp. 528–543, 2020.

[17] C. E. Kasworm, "Adult students: a confusing world in undergraduate higher education," *The Journal of Continuing Higher Education*, vol. 66, no. 2, pp. 77–87, 2018.

[18] R. S. E. Park and J. Scott-Clayton, "The impact of Pell Grant eligibility on community college students' financial aid packages, labor supply, and academic outcomes," *Educational Evaluation and Policy Analysis*, vol. 40, no. 4, pp. 557–585, 2018.

[19] F. Yun, "Reconstruction of primary school students' concept of labor life in we media era," *US-China Education Review*, vol. 10, no. 6, pp. 281–287, 2020.